

STIC-Biotech/ChemLib

68491

Fr m:
Sent:
To:
Subject:

Li, Ruixiang
Monday, June 10, 2002 4:28 PM
STIC-Biotech/ChemLib
Sequence search of Application NO: 09/733,387

Please do a standard search on SEQ ID NOs: 43 and 44 against both the commercial and interference nucleic acid database.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10C01
306-0282

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

RECEIVED
JUN 11 2002
STIC

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/11/02
Date Completed: 6/12/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: 1
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: A
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 20:11:27 ; Search time 1691.49 seconds
(without alignments)
13165.893 Million cell updates/sec

Title: US-09-733-387-43

Perfect score: 1650

Sequence: 1 atggcgagccagggcct.....actcgcacatcagaagaatg 1650

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_esti.*
10: gb_est2.*
11: gb_hic.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	613.4	37.2	779	BF183209	BF183209 601809910
2	443.4	26.9	526	BM193875	BM193875 TCAAP1E43
3	439.4	26.6	736	EG975242	EG975242 602843148
4	399.2	24.2	930	BF788639	BF788639 602105522
5	394.2	23.9	436	BM145287	BM145287 TCAAP1D46
6	351.2	21.3	376	BM145497	BM145497 TCAAP1D64
7	338	20.5	585	BE281756	BE281756 601099701
8	314	19.0	800	BI646880	BI646880 603278628
9	309.2	18.7	518	AA981381	AA981381 vx56d06.r
10	299.8	15.7	338	BE241639	BE241639 TCAAP1E06
11	229.6	13.9	391	UI4114	UI4114 MM014114 ca
12	216.4	13.1	1087	AK014445	AK014445 Mus muscu
13	207.8	12.6	432	BE831720	BE831720 RCO-MT005
14	191.6	11.6	856	BI660803	BI660803 603303847
15	168.4	10.2	329	BM484584	BM484584 538496 MA
16	166.6	10.1	248	BM146119	BM146119 TCAAP1E14
17	151.8	9.2	3835	BC013207	BC013207 Homo sapi

18	145.8	8.8	1038	10	BM450054
19	143.6	8.7	953	10	BF164505
20	142.2	8.6	519	12	AQ171270
21	140.2	8.5	847	10	BI257601
22	140.2	8.5	567	12	AQ607309
23	139.8	8.5	816	10	BG760961
24	130	7.9	446	9	BB860353
25	119.4	7.2	823	10	BF345756
26	117	7.1	560	10	BI066761
27	116.2	7.0	677	10	BG818779
28	115.6	7.0	453	9	BB860092
29	114.2	6.9	537	10	BE754100
30	113.2	6.9	736	10	BI661416
31	111.4	6.8	431	9	AW592568
32	111.4	6.8	580	9	AI807052
33	110.2	6.7	790	10	BI691050
34	109.4	6.6	756	10	BI091842
35	108	6.5	600	10	BG804206
36	105.8	6.4	756	10	BG770361
37	105.8	6.4	934	10	BG032841
38	104.2	6.3	523	10	BG004188
39	104.2	6.3	883	10	BI755503
40	102.4	6.2	794	10	EG914136
41	99	6.0	659	10	BF783449
42	98.6	6.0	893	10	BF784683
43	97.8	5.9	753	10	BF584015
44	96.2	5.8	946	10	BF346638
45	95.4	5.8	733	10	BG833698

ALIGNMENTS

RESULT 1

BF183209 779 bp mRNA linear EST 31-OCT-2000
LOCUS 601809910F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:404042 5',
DEFINITION mRNA sequence.
ACCESSION BF183209.1 GI:11061396
VERSION BF183209.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 779)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM858 row: d column: 03
High quality sequence stop: 701.
Location/Qualifiers
1. 779
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:404042"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,

FEATURES source

1038 10 BM450054
953 10 BF164505
519 12 AQ171270
847 10 BI257601
567 12 AQ607309
816 10 BG760961
446 9 BB860353
823 10 BF345756
560 10 BI066761
677 10 BG818779
453 9 BB860092
537 10 BE754100
736 10 BI661416
431 9 AW592568
580 9 AI807052
790 10 BI691050
756 10 BI091842
600 10 BG804206
756 10 BG770361
934 10 BG032841
523 10 BG004188
883 10 BI755503
794 10 EG914136
659 10 BF783449
893 10 BF784683
753 10 BF584015
946 10 BF346638
733 10 BG833698

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 136 a 229 c 218 g 196 t
ORIGIN

```

Query Match          37.2%; Score 613.4; DB 10; Length 779;
Best Local Similarity 97.5%; Pred. No. 2.4e-127;
Matches 676; Conservative 0; Mismatches 11; Indels 6; Gaps 5;

QY 662 ctaaaggaccactgagactgtctctgagggctgtccacgaggtcagacctgagg 721
DB 2 CTAAGAGGACCACTGGAGACTGGTCTGTGAGGGCTGTCCACGGAGGTACAGACTGAGG 61
QY 722 ggaacctgtgtctgtgtgaccacactgaccttttgcctctgtctgtgacccacctgg 781
DB 62 GGACCGTGTGTGTGTGACCACTGACCTTTTCGCCCTGTCTCTGAGACCCACCTGG 121
QY 782 accagtccacggtgcatactccacacgcatctccagggcggtgtgggtctccatga 841
DB 122 ACCATCCACGGTGCATATCTTCACACGCATCTCCAGCGGGCTGTGGGFTCTCCATGA 181
QY 842 tctctggcctccaccattatttattgctcttctgaggttttccgaggagaggttca 901
DB 182 TCTTCTGGCCTTCACCATTTATCTTTATGCTTTCTGAGGCTTTCCCGGAGAGGTTC 241
QY 902 agtcagaagatgccccaaagatccacgtggtggtggcagctgttctctcgaatc 961
DB 242 AGTCAGAAGATGCCCAAGATCCACGTGG-CCTGGGTGGCAGCGCTGTCTCTCTGAATC 300
QY 962 tggcctcttggtaagtggagtgactcaaaagggtctgactgctgctgctgctggccc 1021
DB 301 TGGCCTTCTTGGTCAATGTGGGAGTGGCTCAAGGGGTCTGATGCTGCTGCTGGGCC 360
QY 1022 ggggggctgtctccactacttctgctgtgcttccactcgtgagggccttgaagcct 1081
DB 361 GGGGGGCTGTCTTCCACTACTTCTGCTGTCTGCTTCCACTGGATGGCCTTGAAGCT 420
QY 1082 tccaccttaactgctgctgaggtcttcaacactaatttgggcaactacttctcga 1141
DB 421 TCCACCTCTACTCTGCTGCTGAGGGTCTTCAACACCTACTTGGGCACTACTTCTCTGA 480
QY 1142 agctgagcctggtgggctgggctgcccgcctgagtgctcagcactggagtgcca 1201
DB 481 AGCTGAGCCTGTGGGTGGGCTGGGCTGCCCGCTGATGCTCATCGCACTGGAGTGCCA 540
QY 1202 acagctacggccttacaccatcctgtagaggagacccgacactctctgagctatgct 1261
DB 541 ACAG--TACGGCTCTACACCATCGTGATAGGGAGAAACGCACTCTCTGGAGCTATGCT 598
QY 1262 ggtt-cggtgaagggaacaaccatgacgcccctctatatcaccgtccacggctacttctc 1320
DB 599 GGTTCCTGTGAAGGGAGCAAAATGTAAGCGCTCTATATACCGCTCCACGG-TAATTCCTC 657
QY 1321 atcactctcttctgtgcatggtgctgccc 1353
DB 658 ATCA-CTTCTCTTTGGCATGGTGGTCTCTGCCC 689

```

RESULT 2

BM193875

LOCUS

DEFINITION TCAPLE1365 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCAAP4365, mRNA
sequence.

ACCESSION BM193875

VERSION BM193875.1 GI:17652472

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 526)

Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
Gunnarathne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
Unpublished (2001)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@tccc.org

Seq primer: M13 primer.

Location/Qualifiers

1..526

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="TCAAP4365"

/clone_lib="Pediatric pre-B cell acute lymphoblastic

Leukemia Baylor-HGSC project-TCBA"

/sex="male"

/tissue_type="leukopheresis"

/cell_type="pre-B cell"

/dev_stage="pediatric 2 years"

/lab_host="DH10B"

/note="vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'GGAGGACTGCGGCGCGGAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGGCTGGATCGCGCGGCAATATATAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pSB vector. Library went through one round of

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T,

Itoh W., Nagaoka S., Sasaki N., Okazaki Y., Muramatsu M,

Schneider C., Hayashizaki Y., High efficiency selection of

full-length cDNA by improved biotinylated cap trapper.,

DNA Res 4: 1, 61-6, Feb 28, 1997)

BASE COUNT 125 a 156 c 152 g 93 t

ORIGIN

Query Match

Best Local Similarity 26.9%; Score 443.4; DB 10; Length 526;

Matches 447; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atggcgagccagggggcctggggcctgctctgctctgctctgctctgctcaggt 60

DB 73 ATGGCGAGCCCGAGGGGCTGGGGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 132

QY 61 caggaaagccacgagggcccaagaaacacacctgctgggagcaacaacatgtacac 120

DB 133 CAGAAAGCCCGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 192

QY 121 atcttcaacttgatgacaggtcttctgcttcaacagtcagcagctcggcgagc 180

DB 193 ATCTTCAACTTGAATGACAAGCTTTGTGCTTCCACCAAGTCAGGCGAGTCGCGG 252

QY 181 tcttgcactgtggaacacttgacagatactggtctaaactacgagggcccatctgatga 240

DB 253 TCTTGCATGTGGAACCTTCAGAGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312

QY 241 gaaggttgacgagaggtgaacacaccttctctgaggttgggttcagaaacctcagc 300

DB 313 GAAGGTTGACGCAAGAGGGAACACCGCTTCTCTGAGGCTTTGGTCTCAGAACCTC 372

QY 301 accaactgacagagacttcttctctctgagccctcaggttccgagggcaggtg 360

DB 373 ACCAACACTGCAGAAGACTTCTATTCTCTCTGAGCGCTCTCAGGTTCCGAGGCG 432

QY 361 atgaagacagaggaagcccccctgacagagtgacgttcgacttcccaagagccttttcgatcc 420


```
|||||
Db 433 ATGAGGAGGACAAAGCCCTGACAGAGTGCACACTTCCCAAGAGCCTTTTCGATCC 492
QY 421 ctgccaggcaacaggtctgtgtcgcgttgccc 453
Db 493 CTGCCAGCACAGGTCTGTGGGCGCGCTGGCC 525

RESULT 3
BG975242 736 bp mRNA linear EST 12-JUN-2001
LOCUS 602843148F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4979170 5',
DEFINITION mRNA sequence.
ACCESSION BG975242
VERSION BG975242.1 GI:14362879
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 736)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10976 row: o column: 11
High quality sequence stop: 736.
FEATURES
source
1. 736
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4979170"
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lothar Hennighausen/Priscilla Furth, NIH
Reference for transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NIH_CGAP
Library."
BASE COUNT 123 a 227 c 180 g 206 t
ORIGIN

Query Match 26.6%; Score 439.4; DB 10; Length 736;
Best Local Similarity 76.4%; Pred. No. 2.8e-88;
Matches 569; Conservative 0; Mismatches 166; Indels 10; Gaps 2;

QY 808 cgcattccccggggcgtgggtctccatgatctctctggcctcaccattattctt 867
Db 1 CGCATCTCCAGGAGGCGGCTCCCATGATCTCTCGGCTTCACCATGGTGTCT 60
QY 868 tatgctcttctgagccttccggagaggttcagtcagaagatgccccaaagatccac 927
Db 61 TATGTTGGCTTCTCTCGCAGAGGTTCAAGTCTGAAGATGCCCTTAAGATCCAC 120
QY 928 gtggcccttgggtggcagccttctcctgaatcgtggccttctgtgtaattgtgggagt 987
Db 121 ATGGCTCTGAGCATCAGCCTGTTCCTCTGAATCTTACCTCTCCGTGATCAATGTGGGAGC 180
QY 988 ggctcaaaagggtctgatgctgctgctgggcccgggggtgtgtcttccactactctg 1047
```

```
|||||
Db 181 AGCTCCCAAGCCCGCCAGCCTCTCTGCTGGTCCGAGCTGCCACTTTTCCACTATTCTG 240
QY 1048 ctctgtgccttoacctgagtgggccttgagccttccacctctacacctgctgcgtgtcagg 1107
Db 241 CTCTGTGCTCTCACTACCTGGTGTGGAGGCTTCCACCTTACCTGCTGGGCGCATCAGA 300
QY 1108 gtcttaaacaccttctcggcactacttctgaagctgagcctggggtgggctgggctg 1167
Db 301 GTCCTCAATACCTACTTTGGACACTATTCTCTGAAGCTGAGCTGCTGGGCTGGGGGCTG 360
QY 1168 cccgccctgagtggtcactcgccactgggagtgcccaacagctacggcctctacacacatc 1227
Db 361 CCTGTCTCTGTAGTTATTGGTGTGGAAGTAGCAACAGTTATGTAGATTACACCATCGC 420
QY 1228 gatagggaagaccacacctctctgagcctgagcctggttccgtgagggagacaacctatc 1287
Db 421 GACCAGGAGAACCCACATCATCTGGAGTTGCTGGTTCAGAAAGAGCC-----T 471
QY 1288 gccctctatatacccgctccacggtactctctcctcactcctctcttggcgtgtgttc 1347
Db 472 GCTCTTTATGCTACTGTCCAGCGGTACTTCTCTGTACCTTCCCTTCTCGGTGTGTGTA 531
QY 1348 ctggccctgtgtgttggaagattctcaccctgtcccgctgctacagcggtcaagagcgg 1407
Db 532 CTGGCTTTGGTGGCTGGAAGATCTTCACTCTGCCAGAGTCACAGCGGGGCAAGGACAG 591
QY 1408 gggagaagaccggaagagtgctcaccctgctggcctctcagcctgtgtgtgtgtgaca 1467
Db 592 GGGCGGAGCTGGAAGTCGGTCTCACCCTGCTGGGCTCTCAGTCTGTGGGCGATGACC 651
QY 1468 tgggggttgccattctcaccctgtggcctctcaccctctcaccctctacatttgcactttc 1527
Db 652 TGGGGGCTGGCTGTCTCACCCTCGCTTGTCTACCATCTAGTCTTCCACCTCTTA 711
QY 1528 aa-ctccttgcaaggtgtcttcac 1551
Db 712 AACCTCTCTACAAGGGCTCTTCATC 736

RESULT 4
BF788639 930 bp mRNA linear EST 12-JAN-2001
LOCUS 602105522F2 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223741
DEFINITION 5', mRNA sequence.
ACCESSION BF788639
VERSION BF788639.1 GI:12093675
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9812 row: k column: 06
High quality sequence stop: 673.
FEATURES
source
1. 930
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4223741"
```

/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 185 a 268 c 253 g 224 t
ORIGIN

Query Match 24.2%; Score 399; DB 10; Length 930;
Best Local Similarity 72.2%; Pred. No. 3.5e-79;
Matches 564; Conservative 0; Mismatches 205; Indels 12; Gaps 3;
QY 236 tgaaggaaggttgagcagaaggttaaacacgccttctcgaagcttggttccagaacc 295
Db 12 TGCTGGAGAAATTCATGGAGACTGTGACATGCTCTTTGTTAAGGATTCATCCAGACA 71
QY 296 tcagcaccacactgcagaagactttattctctgagccctctcaggttccagggc 355
Db 72 TCAGCAGCGATGCTCAGAAAGACTGCT-GTACTCTCTGATGCTTTCCAGATCCCGAGGC 130
QY 356 agtgatgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 415
Db 131 AGGTGATGCAAGGTGAGGAGCGCTGCTGATGGATGCGACTTCCAGAGCGCTGTTG 190
QY 416 gatccctgccagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 475
Db 191 GAGCCCTGCTGCAACAGCTGTCAGTCCGGTGGCCATAACTGTTCTAGACATTTGGT 250
QY 476 cagggaactctctcaagggcccccgcctgcgcctggagagatgacagcgcgtgtgaaca 535
Db 251 CAGGGAATGCTTTAAGGGCCCCAAGCTCCTTGAGGACAAAGGACAGCGTGTGAACA 310
QY 536 atgcctcgttggttggttggttggttggttggttggttggttggttggttggt 595
Db 311 ACCGATGTTGGTTGAGCGTGGCCAGATGACCCCGGCTGTCGGAACCTGTAG 370
QY 596 agatcgtctctctcaccagcagcagcagcagcagcagcagcagcagcagcagcag 655
Db 371 AGATACCTTCTCCCATGAGCGTCAACACCTTAACGTGATCTCTCACTGTATTCTGG 430
QY 656 atgtgactaaagggacacactggagactggtctctgagggctgctccagcaggtcac 715
Db 431 ATATGGCTAAGGAGACT-----GGATTCACCGCTGCTCCAGGTCGCCGGG 480
QY 716 ctgaggggaccgtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 775
Db 481 ATGGGAGAACTGCTGCGGTGGGACACTTGACCTTCTTCGCTTTCGCTGTGAGGCCGA 540
QY 776 ccttgaccagtcacaggtgcatctcctcacacgcatctccagcagcagcagcagcag 835
Db 541 TCTTGGACCTGGCCAGCCACACACTCTCACTCGCATCTCCAGGAGCAGCATCTGCTGG 600
QY 836 ccatgatctctcgtgcttcttccatattcttctgagggcttcccccggag 894
Db 601 CCATGATCTTCTTGGCTTCCACCATGCTGCTCTATGATGCTTCAGGTTCTCTCTGCA 660
QY 895 aggttcaagtcagaagatgcccaagatccacgtgagcctgggtggcagcgtgttcctc 954
Db 661 AGGTTCAAGTCTGAAGATGCCCTTAAGATCCACATGGTTTCTGAGCATCAGCTGTCT 720
QY 955 ctgaattggcctctcttggttcaatgtggggagtggttcaaaaggggtctgatgctgctgc 1014
Db 721 CTGAATCTTACCTTCTGATCACTGTGGGGAGAGCTCCAAAGGCCCCAGCATCTGCTGG 780
QY 1015 t 1015
Db 781 T 781

RESULT 5

BM145287/c

LOCUS BM145287 436 bp mRNA linear EST 30-NOV-2001
DEFINITION TCAAPD4626 Pediatric acute myelogenous leukemia cell (FAB M1)
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP4626, mRNA
sequence.
ACCESSION BM145287
VERSION BM145287.1 GI:17163166
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 436)
AUTHORS Wei, Y., Tsang, Y. T. M., Mei, G., Ku, J. M., Ali-Osman, F. R. Jr.,
Gundratne, P. H., Muzny, D., Bouck, J., Gibbs, R. A., and Margolin, J. F.
Pediatric Leukemia cDNA Sequencing Project (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Seq primer: M13 primer.
FEATURES
Location/Qualifiers
1..436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAAP4626"
/clone_lib="Pediatric acute myelogenous leukemia cell (FAB
M1) Baylor-HGSC project-TCAA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'-GGAGACTCGAGCGCCGCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'-AGAGAGTCGATCCGCGCGCGCAATAATAATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 98 a 124 c 140 g 74 t
ORIGIN
Query Match 23.9%; Score 394.2; DB 10; Length 436;
Best Local Similarity 96.5%; Pred. No. 3.2e-78;
Matches 414; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
QY 1014 ctgggcccgggggctgtctccactactctctctctctctctctctctctctctctctct 1073
Db 429 CTGGCGCGGAGGGGTGTCTCCCGCGCTGTCTGTGTGCTTCACTGGATGGGCTT 370
QY 1074 tgaagccttccacacctctacacctgctcgctgacagggctctcaacacctactctcggcact 1133
Db 369 TGAAGCCTTCCACCTCTTACCTGCTCGCTGTCTCAGGGTCTTCAACACCTACTTTCGGGCACTA 310
QY 1134 ctctctgaagctgagcctgtggtgggctgggctgggctgggctgggctgggctgggctggg 1193
Db 309 CTTCTGAAGTGAAGCTGTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 250
QY 1194 gagtgcccaacagctacggcctctacacacctcgtgtagaggagagacccgacacctctctgga 1253
|||||

LOCUS AA981381 518 bp mRNA linear EST 27-MAY-1998
DEFINITION vx56d06.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
clone IMAGE:1279211 5' similar to TR:000406 000406 SEVEN
TRANSMEMBRANE-DOMAIN RECEPTOR. ; mRNA sequence.
ACCESSION AA981381
VERSION AA981381.1 GI:3159917
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 518)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilton, R. and
Waterston, R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:671011
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 439.
FEATURES
Location/Qualifiers
1..518
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1279211"
/clone_lib="Stratagene mouse macrophage (#937306)"
/tissue_type="macrophage"
/dev_stage="WEHI-3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'";
BASE COUNT 85 a 152 c 131 g 150 t
ORIGIN
Query Match 18.7%; Score 309.2; DB 9; Length 518;
Best Local Similarity 75.9%; Pred. No. 4.1e-59;
Matches 399; Conservative 0; Mismatches 118; Indels 9; Gaps 1;
QY 1050 ctgtgcctcacctggatggccttggaagccttcacacctctacacctctcgtcgtgcagggt 1109
|||||
DB 1 CTGTGCTTCCACCTGGATGGTCTGGAGGCTTCCACCTCTACCTGCTGGCCATCAGAGT 60
QY 1110 cttcaacacactctctggcactactctcctgaagctgagcctgggtggcctgccc 1169
|||||
DB 61 CTTCAATACCTACTTTTGACACATATTTCTTGAAGCTGAGCTGCTGGCCTGGGCTGGCG 120
QY 1170 cgcctgatgtcagcgcactggagctgcaacagctcagcctctacacacatccgtga 1229
|||||
DB 121 TGTCTTGTAGTATTGCTGGAGTAGCAACAGTTATGGAGTTTACACCAATTCGCGA 180
QY 1230 tagggagaacgcacactctctgagctatgctggttcctggaagggaacacactgtacgc 1289
|||||
DB 181 CCAGGAGAACCGACATCACTGGAGTTGCTGCTGCTCCAGAAAGAGCC-----TGC 231
QY 1290 cctctatatcacgcgcgcactctctctcctcctcctcctcctcctctctctctctct 1349
|||||
DB 232 TCATTATGCTACTGTCCACGGCTACTTCCCTGTTTACCTTCTCTCTCGTGGTGGTACT 291

QY 1350 ggccctgggtgtctggaagatcttcacccctgtcccggtgtcacagcggtcaaggagcggg 1409
|||||
DB 292 GGTCTTGGTGGCTGGAGATCTTCACTCTGCCAGCTGTTCACAGCGGCAAGACAGGG 351
QY 1410 gaagaacccggaagaaggtgctcaccctgtggcctcgtcagcctggtgggtgtgacatg 1469
|||||
DB 352 GCGACCTGGAAGTGGGTCTCTACCGTGTGGGCTCTCCAGTGTGGGCGATGACCTG 411
QY 1470 ggggtggccatcttcacccctgtggcctcctcaccgtctacatcttgcacatttcaa 1529
|||||
DB 412 GGGCTGGCTGTCTCTACCCCTCGGCTGTGTACCATCTACGCTCTACCCCTCTTAAA 471
QY 1530 ctcttcgaaggtgtcttcacctctcgtcgtggttcaccacacctcttac 1575
|||||
DB 472 CTCTTACAAAGGCTCTTCATCTTCTGCTGGTTTCATCATCTCTCTAC 517
RESULT 10
BE241639
LOCUS
DEFINITION TCAAP1E0686 Pediatric acute myelogenous leukemia cell (FAB M1)
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP0686, mRNA
Sequence.
ACCESSION BE241639
VERSION BE241639.1 GI:9093362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 338)
AUTHORS Wei, Y., Tsang, Y. T. M., Mei, G., Ku, J. M., Ali-Osman Jr., F. R., Muzny, D.
Bouck, J., Gibbs, R. A. and Margolin, J. F.
TITLE Pediatric Leukemia cDNA Sequencing Project
JOURNAL Unpublished (2000)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@ccc.org
Citation: Carninci, P. and Hayashizaki, Y. High efficiency
full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Seq primer: M13 primer.
FEATURES
Location/Qualifiers
1..338
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAAP0686"
/clone_lib="Pediatric acute myelogenous leukemia cell (FAB
M1) Baylor-HGSC project-TCAA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH108"
/note="Vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTGCGCGCGGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGCTCGGTCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda PSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci, P., Westover, A., Nishiyama, Y., Ohsumi, T.,
Itoh, M., Nagaoka, S., Sasaki, N., Muramatsu, M.,
Schneider, C., Hayashizaki, Y., High efficiency selection of
full-length cDNA by improved blotting cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997")
BASE COUNT 83 a 97 c 101 g 57 t

```

Query Match      13.9%; Score 229.6; DB 10; Length 391;
Best Local Similarity 75.5%; Pred. No. 2.7e-41;
Matches 302; Conservative 0; Mismatches 89; Indels 9; Gaps 1;

QY 972 ggtcaatgtggggagtggtcctcaagggtctgaatgctgctgagccggggggctgt 1031
      |||||
Db 391 GATCAATGTGGGAGCAGCTCCCAAGCCCCCAGCCTCTGCTGGGTCCGAGCTGCCAT 332
      |||||

QY .1032 cttcaactactctctgtctgtgccttcacbtggatgggacctgaagccttcacactcta 1091
      |||||

```

11042159	3 (sites)	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
		Genome Res. 10 (11), 1757-1771 (2000)
20530913		
11076861	4 (sites)	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
		Functional annotation of a full-length mouse cDNA collection
		Nature 409, 685-690 (2001)
5 (bases 1 to 1087)		
		Adachi, T., Aizawa, K., Akabira, S., Akimura, T., Akao, H., Akashi, M.,

QY	1352	ccttggtggtcgaagaatcttcacccctgtcccggtctacaagcgcgaaggcggggga	1411
DG	61	CTTTGGTGGCGTGGAAGATCTCACTTCGCCAGTGTACAGCGGGCAAAGGACAGGGGC	120
QY	1412	agaaccggaagaagtgtctcacctcgtgggcctctcgagcctggtgggtgtgacatggg	1471
DG	121	CGACCTGGAAGTCGGTCTCACCGTGTGGCCTCTCCAGTCTGTTGGCATGACCTGGG	180
QY	1472	ggttggccatcttcacccccttgggcctctccaccgtctacatatottgcactttccaact	1531
DG	181	GGCTGGCTGTTCTACGCCCGCTGGCTTGCTACCATCTACGCTCTTCAACCTCTAAACT	240
QY	1532	cettgcgaagtgattcatctgctgtgttcaacctcttttaacctccccagtcagagca	1591
DG	241	CFTCTAAGAAGGCTCTTCATCTTGCTGGTGTATCATCTCTACTTCCTCTACTCAGAGCA	300
QY	1592	ccacagtcctctctc---tactgcaagattggaccaggcccaactcccgaatctcaaagaat	1648
DG	301	CCACAGCCTCTCTCTCCGGCACC GGCGGTGGACAGCCACCTCCCGTGTCCAGGAGT	360
QY	1649	ag	1650
DG	361	AG	362

please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Db		361 AG 362	
RESULT	13		
BE8311720/c			
LOCUS	BE8311720	432 bp mRNA linear EST 22-SEP-2000	
DEFINITION	RC0-MT0059-200600-021-a08 MT0059 Homo sapiens cDNA, mRNA sequence.		
ACCESSION	BE8311720		
VERSION	BE8311720.1	GI:10264098	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 432)		

```

cna: bam: /mnt/scratch/cna/chr10_100000000-100000000.bam
location/Qualifiers
    l.: 1087
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="MGD:MGI:1903138"
    /db_xref="taxon:10090"
    /clone="383042G05"
    /sex="female"
    /tissue_type="placenta and extra embryonic tissue"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="18 days pregnant adult"

```

744. . 1070
/gene- 9p13/

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimposon@ludwig.org.br

```

padding=
/codon_start=1
/protein_id="BAB29355.1"
/db_xref="GI:12852301"
/translation="MGPKYKCRSAGLSPCKPWPSPQQCYLPCCYGDRLSSCG
PPIHDETRTEACSIYIPWRGKREKKOLFRRLYSGTALPTISTQPEHQAWLCOAN
NCHTFK"
252 a 306 c 250 q 279 t

```

NUMBER			
252 a	306 c	250 g	279 t

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="WT0059"
/dev_stage="Adult"

```

SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 107 a 108 c 128 g 89 t

ORIGIN

Query Match 12.6%; Score 207.8; DB 10; Length 432;
Best Local Similarity 99.1%; Pred. No. 2.2e-36;
Matches 209; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1440 ggccctcgcagcctgggtgacatggggtggcctatccaccccggtggcct 1499
|||||
Db 432 GGGCTCTCGAGCCTGGTGGGTGACATGGGGGTGGCCATCTTACCCCGTGGCCCT 373
|||||

QY 1500 ctcacccgtctacatctttgcaactcttcaactcctgcaagggtgttctatctgtctg 1559
|||||
Db 372 CTCACCCGTCTAAATCTTTGCACTCTTCAAGGTGCTTCACTCTGCTGCTG 313
|||||

QY 1560 gttcaccatcttaccctcccaagtcagagaccacacagtcctctctactgcaagatt 1619
|||||
Db 312 GTTCAACATCTTTACCTCCCAAGTCAGACACCACTCTCTCTCTACTGCAAGATT 253
|||||

QY 1620 gacacagcccccactccgcctatcgaagaatag 1650
|||||
Db 252 GGACCAGGCCCCACTCCGCATCTCAAGAATAG 222
|||||

RESULT 14

B1660803 B1660803 856 bp mRNA linear EST 12-SEP-2001
LOCUS 603303847F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5349707 5',
DEFINITION mRNA sequence.

ACCESSION B1660803.1 GI:15575039

VERSION EST.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.

1 (bases 1 to 856)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11888 row: n column: 12

High quality sequence stop: 811.

Location/Qualifiers

1..856

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:5349707"

/clone_lib="NIH_CGAP_Mam4"

/tissue_type="tumor, gross tissue"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert 2.5 kb. Library constructed by Life

Technologies, catalog # 12018-016. Investigators providing

samples: Lothar Hennighausen/Priscilla Furth, NIH

Reference for transgenic model: Li et al., Cell Growth and

Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP

Library."

BASE COUNT 177 a 241 c 246 g 192 t

ORIGIN

Query Match 11.6%; Score 191.6; DB 10; Length 856;
Best Local Similarity 77.2%; Pred. No. 1.3e-32;
Matches 233; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 339 ctctcaggtctccgagcaggtgatgaagacgacgacacagccccctgacagagtgcagct 398
|||||
Db 223 CTTACAGATCCCGAGCAGGTGATGACGGTGAGGACGAGCCTGCTGATGGAGTCCGACT 282
|||||

QY 399 tcccaagagccttttctgatccctgcagagcaacaggtctgtggtccgcttggccgtcac 458
|||||
Db 283 TCCCAAGAGCCTGTTTGGAGCCCTGCTGGCAACAGGTCTGCAGTCCCGTTGGCCATAAC 342
|||||

QY 459 cattctggacattgggtccaggggactctctcaaggggcccccggtcggcgtggagatgg 518
|||||
Db 343 TGTTCATACATTTGGTGCAGGAATGCTTTAAGGGCCCCAAGCTCCTTGAGGACAAGGG 402
|||||

QY 519 cagcggcgtgtgaacaacgcctgctggtttgagttggtgggacacaaatgcattgcacca 578
|||||
Db 403 CAGCAGCCTTTGAACAACCGCATGTTGGTGTGGCCAGATGCACGCCACCGG 462
|||||

QY 579 gctggctgagcctctggagatgcttctctcaccagggacccgccccctaacatgacct 638
|||||
Db 463 GCTGTCGGAACCTGTAGAGATCACCTTCTCCCATGAGCGTCAACCACTGCAATGATCCT 522
|||||

QY 639 ca 640

Db 523 GA 524

RESULT 15

LOCUS

BM484584

DEFINITION

538496 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.

ACCESSION

BM484584

VERSION

BM484584.1 GI:18534912

KEYWORDS

EST.

SOURCE

pig.

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 329)

Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,

Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.

and Keeler, J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 782 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCCAGTCAGCAG

Plate: 4 row: P column: 13

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..329

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARC 2P1G"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from testis, ovary,

BASE COUNT	90 a	88 c	86 g	65 t
ORIGIN	endometrium, hypothalamus, pituitary, and placenta."			

Query Match	10.2%	Score 168.4	DB 10	Length 329
Best Local Similarity	74.8%	Pred. No. 1.4e-71		
Matches 211	Conservative 0	Mismatches 27	Indels 0	Gaps 0
Qy 108	caacatgtacgacatcttcaacttgaatgacaaggctttgtcttcaccaagtgcaggca	167		
Db 47	CAACCAGTACGATAACTTCGGCTTGAAGACACGGCTAAGTCTTCGCAAGTGTACACA	106		
Qy 168	gtcgggcacgactcctcaatgtggaaccttcagagatactgctaaactacagggc	227		
Db 107	GTGGGGACAGGATCCTGCAATCTGGGAATCTCAGAGATCTGGCTGAGCTATGAGT	166		
Qy 228	ccaatctgtatgaagggaagtttgacgcagaaggtggaacacgcttttctgaaggtttggt	287		
Db 167	CCATCTGTGTGGAGACAAGTCCAAAGGATACGGTGAACATGCTCTTTTGAAGACTCTGT	226		
Qy 288	cgagaaccttcagcaccaactcgagaagactctattctctctgagccctctcaggt	347		
Db 227	TCTGAACATCAGCACCAACATCTCGAAGACCTGCATCTTCTCCCTGGCGCCTCTCAGAT	286		
Qy 348	tcocgaggcagggtgatgaagcagcaggacaagcccccttgacag	389		
Db 287	TCCAAGGCAAGGTGACGGAGGACGACACACGACCCCGACCG	328		

Search completed: June 11, 2002, 20:49:55
Job time: 2308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 20:16:47 ; Search time 2132.3 seconds
(without alignments)
16193.212 Million cell updates/sec

Title: US-09-733-387-43
Perfect score: 1650
Sequence: 1 atggcgagcccgaggccgt.....actcgcatctcaagaatag 1650

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_btg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description

1	1650	100.0	1650	6	AX167301	AX167301 Sequence
2	1543	93.5	1566	6	AX167281	AX167281 Sequence
3	1454.6	88.2	1527	6	AX167261	AX167261 Sequence
4	1441	87.3	1458	6	AX167291	AX167291 Sequence
5	1430.6	86.7	1515	6	AX167271	AX167271 Sequence
6	1161	70.4	1203	6	AX167305	AX167305 Sequence
7	1054	63.9	1119	6	AX167285	AX167285 Sequence
8	1020	61.8	1020	6	AX167303	AX167303 Sequence
9	965.6	58.5	1080	6	AX167265	AX167265 Sequence
10	952	57.7	1011	6	AX167295	AX167295 Sequence
11	941.6	57.1	1068	6	AX167275	AX167275 Sequence
12	913	55.3	936	6	AX167283	AX167283 Sequence
13	824.6	50.0	897	6	AX167263	AX167263 Sequence
14	811	49.2	828	6	AX167293	AX167293 Sequence
15	800.6	48.5	885	6	AX167273	AX167273 Sequence
16	772.6	45.8	825	6	AX167307	AX167307 Sequence
17	747.8	45.3	2127	10	AF249738	AF249738 Mus muscu
18	665.6	40.3	741	6	AX167287	AX167287 Sequence
19	628.4	38.1	678	6	AX167259	AX167259 Sequence
20	577.2	35.0	702	6	AX167267	AX167267 Sequence
21	563.6	34.2	633	6	AX167297	AX167297 Sequence
22	553.2	33.5	690	6	AX167277	AX167277 Sequence
23	490	29.7	4036	6	AX167311	AX167311 Sequence
24	394	23.9	612	6	AX167309	AX167309 Sequence
25	287.4	17.4	78267	2	AC022585	AC022585 Homo sapi
26	287.4	17.4	157765	2	AC018552	AC018552 Homo sapi
27	287	17.4	528	6	AX167289	AX167289 Sequence
28	266.6	16.2	603	6	AX147806	AX147806 Sequence
29	207.2	12.6	161229	2	AC106681	AC106681 Rattus no
30	198.6	12.0	489	6	AX167269	AX167269 Sequence
31	185	11.2	420	6	AX167299	AX167299 Sequence
32	174.6	10.6	477	6	AX167279	AX167279 Sequence
33	161	9.8	3360	6	AX342667	AX342667 Sequence
34	159.4	9.7	1425	6	AX108481	AX108481 Sequence
35	159.4	9.7	1587	6	AX108469	AX108469 Sequence
36	159.4	9.7	3711	9	BC008770	BC008770 Homo sapi
37	159.4	9.7	3711	9	BC008770	BC008770 Homo sapi
38	156.2	9.5	2082	6	AX029423	AX029423 Sequence
39	156.2	9.5	2822	6	AX298003	AX298003 Sequence
40	156.2	9.5	2822	9	AF106858	AF106858 Homo sapi
41	156.2	9.5	2834	6	AX029424	AX029424 Sequence
42	155	9.4	3686	6	AX061726	AX061726 Sequence
43	155	9.4	3845	6	E54092	E54092 Novel gene
44	155	9.4	3845	9	HSA011001	AJ011001 Homo sapi
45	153.4	9.3	3774	6	AX136371	AX136371 Sequence

ALIGNMENTS

RESULT 1	AX167301	AX167301	Sequence 43 from Patent WO0142287.	1650 bp	DNA	linear	PAT 03-JUL-2001
AX167301	LOCUS	AX167301	Sequence 43 from Patent WO0142287.				
	DEFINITION	AX167301	Sequence 43 from Patent WO0142287.				
	ACCESSION	AX167301	Sequence 43 from Patent WO0142287.				
	VERSION	AX167301.1	GI:14596745				
	KEYWORDS						
	SOURCE	human.					
	ORGANISM	Homo sapiens					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	REFERENCE	1 (bases 1 to 1650)					
	AUTHORS	Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B. and Sands,A.T.					
	TITLE	Novel human membrane proteins and polynucleotides encoding the same					
	JOURNAL	Patent: WO 0142287-A 43 14-JUN-2001;					
		Lexicon Genetics Incorporated (US)					
	FEATURES	Location/Qualifiers					
	source	1..1650					
		/organism="Homo sapiens"					
		/db_xref="taxon:9606"					
	BASE COUNT	313 a	507 c	447 g	383 t		
	ORIGIN						

[illegible]

QY	1021	cg9ggggctgtcttcacactacttcctgctgtgcttccacctggatggggccttgaagcc	1080
Db	1021	CGGGGGCTGTCTTCACACTACTTCCCTGCTGTGCTTACCTTGATGGCCCTTGAAGCC	1080
QY	1081	ttcacacctacactgctgcgtgtcagggtttcaaacactactctcggaactactctctg	1140
Db	1081	TTCCACCTCTACCTGTCTGCTGTACAGGGTCTTCAACACCTACTTCGGGCACTACTTCCTG	1140
QY	1141	aagctgagcctggtgggctggggcctgcccgcctgatggtcatcgacgtcgaagtgcc	1200
Db	1141	AAGCTGAGCCTGTGGGCTGGGGCTGCCGCCCTGATGGTCAATCGCACTGGGAGTGCC	1200
QY	1201	aacagctacggcctctacaccatccctgtagaggagaacggcacctctctggagctatgc	1260
Db	1201	AACAGCTACGGCCTCTACACCATCCGTGATAGGAGAAACGGCACCTCTCTGGAGCTATGC	1260
QY	1261	tggttccgtgaagggaacaccatgtacgcccctctatatcacggtccacggctacttcctc	1320
Db	1261	TGGTTCCGTGAAGGACACCATGTAGGCCCTCTATATCACCGTCCACGGCTACTTCCCTC	1320
QY	1321	atcaacttcctctttggcatggtggtccctggcctggtgctggaagatcttcaaccctg	1380
Db	1321	ATCACCTTCTCTTTGGCATGTTGGTCTTCCCTGGCCCTGGTGTCTGGAAGATCTTCACCCCTG	1380
QY	1381	tccctggtcacagcgtcaagagcgggggaagaaacggaaaggtgctcaacctgctg	1440
Db	1381	TCCCGTGTCTACAGCGGTCAAGAGCGGGGGAAGAACCGGAAGAGGTGCTCACCCCTGCTG	1440
QY	1441	ggcctctcagcctgggtgggtgtgacatgggggttggccatcttcaccccgctgggcctc	1500
Db	1441	GGCCTCTCAGCGCTGGTGGGTGTGACATGGGGTTGGCCATCTTCACCCCGTTGGGCCTC	1500
QY	1501	tcacacgtctacatcttgcactttcaactccttgcaaggtgttattcatctgctgctg	1560
Db	1501	TCACCGTCTACATCTTTGCATCTTTCAACTCTCTTGCAAGGTGTCTTCACTGCTGCTGG	1560
QY	1561	ttcacacctcttaacctcccaagtcagagcacacagtcctctctactgcaagattg	1620
Db	1561	TTCAACATCTTTAGCTCCCAAGTCAGAGCACACAGTCTCTCTCTCTACTGCAAGATTG	1620
QY	1621	gacagggccactccgactctcaagaatag	1650
Db	1621	GACcAGGcCCACTCCGcATCTCAAGAAATAG	1650
RESULT	2		
AX167281			
LOCUS	AX167281	1566 bp	DNA
DEFINITION	Sequence 23 from Patent WO0142287.		linear
ACCESSION	AX167281		
VERSION	AX167281.1	GI:14596735	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1566)		
JOURNAL	Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.		
FEATURES	Novel human membrane proteins and polynucleotides encoding the same		
source	Patent: WO 0142287-A 23 14-JUN-2001;		
	Lexicon Genetics Incorporated (US)		
	Location/Qualifiers		
	1..1566		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	294 a	475 c	439 g
ORIGIN	358 t		

Query Match 93.5%; Score 1543; DB 6; Length 1566;
Best Local Similarity 100.0%; pred. No. 9.4e-308;

[illegible]

Qy	1081	ttcacctctacctgctgcgtgctgaagggtcttcaacacacctacttcgggcactactctcgt	1140
Db <td>1081<th>TTCCACCTCTACCTGCTCGCTGTCAGGCTCTTCAACACCTACTTTCGGGCACTACTTCCTG</th><th>1140</th></td>	1081 <th>TTCCACCTCTACCTGCTCGCTGTCAGGCTCTTCAACACCTACTTTCGGGCACTACTTCCTG</th> <th>1140</th>	TTCCACCTCTACCTGCTCGCTGTCAGGCTCTTCAACACCTACTTTCGGGCACTACTTCCTG	1140
Qy <td>1141<th>aagctgaagcctggtgggtgggctgccgcctctgatgtcatcggcactggagtgcc</th><th>1200</th></td>	1141 <th>aagctgaagcctggtgggtgggctgccgcctctgatgtcatcggcactggagtgcc</th> <th>1200</th>	aagctgaagcctggtgggtgggctgccgcctctgatgtcatcggcactggagtgcc	1200
Db <td>1141<th>AAGCTGAGCCTGCTGGCTGGGCTGCCGCCCTGATGTCATCGGCACCTGGAGTGCC</th><th>1200</th></td>	1141 <th>AAGCTGAGCCTGCTGGCTGGGCTGCCGCCCTGATGTCATCGGCACCTGGAGTGCC</th> <th>1200</th>	AAGCTGAGCCTGCTGGCTGGGCTGCCGCCCTGATGTCATCGGCACCTGGAGTGCC	1200
Qy <td>1201<th>aacagtataggcctctacaccatccgtgatagggaagaccctctcttgagctatgc</th><th>1260</th></td>	1201 <th>aacagtataggcctctacaccatccgtgatagggaagaccctctcttgagctatgc</th> <th>1260</th>	aacagtataggcctctacaccatccgtgatagggaagaccctctcttgagctatgc	1260
Db <td>1201<th>AACAGTACGGCCTCTACACCATCGTGTATAGGAGAACCCACCTCTCTGGAGCTATGC</th><th>1260</th></td>	1201 <th>AACAGTACGGCCTCTACACCATCGTGTATAGGAGAACCCACCTCTCTGGAGCTATGC</th> <th>1260</th>	AACAGTACGGCCTCTACACCATCGTGTATAGGAGAACCCACCTCTCTGGAGCTATGC	1260
Qy <td>1361<th>tgtctcgtgaaggacacacatgtacccctctatatcccgctccacggctacttcctc</th><th>1320</th></td>	1361 <th>tgtctcgtgaaggacacacatgtacccctctatatcccgctccacggctacttcctc</th> <th>1320</th>	tgtctcgtgaaggacacacatgtacccctctatatcccgctccacggctacttcctc	1320
Db <td>1361<th>TGTTTCGGTGAAGGACACCATGTACGCCCTCTATATCACCGTCCACGGCTACTTCCTC</th><th>1320</th></td>	1361 <th>TGTTTCGGTGAAGGACACCATGTACGCCCTCTATATCACCGTCCACGGCTACTTCCTC</th> <th>1320</th>	TGTTTCGGTGAAGGACACCATGTACGCCCTCTATATCACCGTCCACGGCTACTTCCTC	1320
Qy <td>1321<th>atcacctctcttggcatggtggtcctggcctggtgctggaagatcttcaacctg</th><th>1380</th></td>	1321 <th>atcacctctcttggcatggtggtcctggcctggtgctggaagatcttcaacctg</th> <th>1380</th>	atcacctctcttggcatggtggtcctggcctggtgctggaagatcttcaacctg	1380
Db <td>1321<th>ATCACCTTCTCTTTGGCATGCTGCTGCGCCTGGTGTCTTGAAGATCTTTCACCTG</th><th>1380</th></td>	1321 <th>ATCACCTTCTCTTTGGCATGCTGCTGCGCCTGGTGTCTTGAAGATCTTTCACCTG</th> <th>1380</th>	ATCACCTTCTCTTTGGCATGCTGCTGCGCCTGGTGTCTTGAAGATCTTTCACCTG	1380
Qy <td>1381<th>tccgtgctacagcgtcaagagcgggggaagaccggaagagtgctcaccctctgctg</th><th>1440</th></td>	1381 <th>tccgtgctacagcgtcaagagcgggggaagaccggaagagtgctcaccctctgctg</th> <th>1440</th>	tccgtgctacagcgtcaagagcgggggaagaccggaagagtgctcaccctctgctg	1440
Db <td>1381<th>TCCGTGCTACAGCGTCAAGGAGCGGGGGAAGAACCGGAAGAGTGCTCACCTGCTG</th><th>1440</th></td>	1381 <th>TCCGTGCTACAGCGTCAAGGAGCGGGGGAAGAACCGGAAGAGTGCTCACCTGCTG</th> <th>1440</th>	TCCGTGCTACAGCGTCAAGGAGCGGGGGAAGAACCGGAAGAGTGCTCACCTGCTG	1440
Qy <td>1441<th>ggcctctcagcctggtgggtgacatgggggtggccatcttcaacccttggggcctc</th><th>1500</th></td>	1441 <th>ggcctctcagcctggtgggtgacatgggggtggccatcttcaacccttggggcctc</th> <th>1500</th>	ggcctctcagcctggtgggtgacatgggggtggccatcttcaacccttggggcctc	1500
Db <td>1441<th>GGCCTCTCAGCCTGCTGGGTGTGACATGGGGTGGCCATCTTACCCCGTTGGGCTC</th><th>1500</th></td>	1441 <th>GGCCTCTCAGCCTGCTGGGTGTGACATGGGGTGGCCATCTTACCCCGTTGGGCTC</th> <th>1500</th>	GGCCTCTCAGCCTGCTGGGTGTGACATGGGGTGGCCATCTTACCCCGTTGGGCTC	1500
Qy <td>1501<th>tccacgctctacatcttgcacttttcaactccttgcagggtg</th><th>1543</th></td>	1501 <th>tccacgctctacatcttgcacttttcaactccttgcagggtg</th> <th>1543</th>	tccacgctctacatcttgcacttttcaactccttgcagggtg	1543
Db <td>1501<th>TCCACCGCTACATCTTGCACCTTTTCAACTCTTTCGAAGGTG</th><th>1543</th></td>	1501 <th>TCCACCGCTACATCTTGCACCTTTTCAACTCTTTCGAAGGTG</th> <th>1543</th>	TCCACCGCTACATCTTGCACCTTTTCAACTCTTTCGAAGGTG	1543
RESULT	3		
AX167261			
LOCUS	AX167261	1527 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 3 from Patent WO0142287.		
ACCESSION	AX167261		
VERSION	AX167261.1	GI:14596725	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Donoho, G., Scoville, J., Turner, C.A., Friedrich, G., Zambrowicz, B. and Sands, A.T.		
TITLE	Novel human membrane proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0142287-A 3 14-JUN-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1527		
BASE COUNT	295 a 459 c 426 g 347 t		
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"		
Query Match	88.2%;	Score 1454.6;	DB 6; Length 1527;
Best Local Similarity	99.7%;	Pred. No. 1.5e-249;	
Matches 1457;	Conservative	0; Mismatches	4; Indels 0; Gaps 0;
Qy	1	atggcgagccagggcctggggcctctctctgtctctctgtctctgtctcccgacctcaggt	60
Db	1	ATGGCGAGCCCGCCAGGGGCTGGGGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	60
Qy	61	caggaaagccaccgaaaggccgaagaacacctctcctggggagcaacaacatgtacgac	120
Db	61	CAGAAAAGCCCGCCAGGAGGCGCAAGAACACCTGCTCTGGGAGCAACACATGTACGAC	120
Qy	121	atcttcaacttgaatgacaaggctttgtgtcttccaaagtgcaggcagtcggtcgagcgac	180

Db	121	ATCTTCAACTTGAATGACAAAGGCTTTGTGCTTCCACCAAGTGCAGGCACTGCGGGCAGCGAC	180
Qy	181	tcctgcaatggaataactgcagagatactggtctaaactacagagcccatctgatgaag	240
Db	181	TCTGTCAATCTGGAATAACTTGCAGAGATACTGGCTAACTACGAGGCCCATCTGATGAAG	240
Qy	241	gaaggtttgacgcagaagtgaacaagcctttccctgaagcctttggtccagaacctaagc	300
Db	241	GAGGTTTGACGCAGAAAGGTGAACAGCCCTTCTCTGAAGGCTTTGGTCCAGAAGCTTCAGC	300
Qy	301	accaaactgcagaagacttctattctctctggagccctctcaggtttccgaggcaggtg	360
Db	301	ACCAAACTTCAGAAAGACTTCTATTCTCTCTGAGAGCCCTCTCAGGTTCCGAGGCAAGTG	360
Qy	361	atgaagacgagacagaagccctcaacagatgcagacttccaagagccttttcgatcc	420
Db	361	ATGAGGACGAGACAAAGCCCTGACAGATGCGACATTCCCAAGAGCCTTTTCGATGCC	420
Qy	421	ctgccaggcaacaggctgtggtccggtctggcgctcaaccattctggacattggtccaggg	480
Db	421	CTGCCAGGCAACAGGCTGTGTGTGCTGCGTTGGCGCTCACCAATTCTGGACAATTGGTCCAGG	480
Qy	481	actctctcaaggcccccggctcgccctggagatggcagcgcggtgtgaacaactgc	540
Db	481	ACTCTCTCAAGGGCCCCGGCTGGCCTGGGAGATGSCAGGGGCGCTTTGAAACAATGCG	540
Qy	541	ctggtgggttgagtgggacaatgcatgtcaaccaagctggtgagcctctggagatc	600
Db	541	CTGGTGGGTTGAGTGTGGGACAAATGCAATGCAACAGCTGGCTTGAGCCCTCTGGAGATC	600
Qy	601	gtctctctcaacagcgacgccccctaaactgaacctcaacctgtattctcggaatg	660
Db	601	GTCTTCTCTCACGAGCGCCCGCCCTAACATGACCTCACCTGTGTATTCTGGGATGTG	660
Qy	661	actaaaggacaactggagactggtctctgagggtcgtccaagggctcagacactgag	720
Db	661	ACTAAAGGGACCACTGGAGACTGGTCTTCTGAGGGCTGCTCCACGAGGTCAGACCTTGAG	720
Qy	721	gggaccgtgtgctgctgtgaccacctgaccttttctgcctgctcctgagaccacactg	780
Db	721	GGGACGCTGTGCTGTGACCACTGACCTTTTTCGCCCTGCTCTGAGACCCACCTTG	780
Qy	781	gaccagtccaagtgatatactccaacgcatctccaagcgggctgtggggctctcaatg	840
Db	781	GACCAGTCCACGGTGATATCTCTACACGATCTCCCAAGCGGGCTGTGGGGTCTCCATG	840
Qy	841	atctctggccttcaccattattcttatgaccttctgaggtcttcccgaggaggttc	900
Db	841	ATCTCTCGGCTTCACCATTAATCTTTATGCTTTCTGAGGCTTTCCCGGAGAGGTTG	900
Qy	901	aagtcagaagatgccccaaagatccaagtcagtgccctgggtggcagcctgttctcctaag	960
Db	901	AAGTCAGAGATGCCCCAAGATCCACGTGGGCCCTGGGTGGCAGGCTGTCTCTGAAT	960
Qy	961	ctggcctcttggtcaatgtygggagtggtctcaaggggctgtagtgcctgctgagggc	1020
Db	961	CTGGCCTCTTGTGTCAAATGTGGGAGTGGCTCAAAAGGGTCTGATGTCGCTGTGGGCC	1020
Qy	1021	cgggggctgtctccaactactctgctgtgacctcaactgagtgagccttgaagcc	1080
Db	1021	CGGGGGGCTGTCTTCACATCTCTGCTCTGTGCTTCACTGATGGAGGGCTTTGAAGCC	1080
Qy	1081	ttccacctctacctgctgctgcaggggtcttcaacacctacttccgggacactactctg	1140
Db	1081	TTCCACCTCTACCTGTCTGCTGTCAAGGGTCTTCAACACCTACTTCGGGCACTACTTCTG	1140
Qy	1141	aagctgagcctggtgagctgggacctccccacctgagtgatgcgaactggagtgcc	1200
Db	1141	AAGCTGAGCCTGGTGGGCTTGGGGCTTCCCGCCCTGATGGGTATCGGCACCTGGGAGTGCC	1200
Qy	1201	aacagctacggcctctacacctccgctgataggggagaaccgcacctctctggagctatgc	1260
Db	1201	AACAGCTAGGCGCTTACACCATCGCTGATAGGAGAACCCGACCTCTCTGGAGCTATGC	1260

[illegible]

Db	601	GTCTTCTCTCACCAGCAGCGCCCGCCCTAAACTGACCCCTCACTGTGTATTCTTGGGATGTG	660
Qy	661	actaaaggaggaactgagactggtctctctgagggtcgtctccacggagggtcagacactgag	720
Db	661	CTAAAGGGACCACTCTGAGACTGGTCTTCTGAGGCGTCTCTCCACGGAGGTCAGACCTGAG	720
Qy	721	gggacagtgctgctgtgtgacacactgacctttttcgccctgtctctctgaaacacacttg	780
Db	721	GGGACCGTGTCTGTCTGACCACTGACCTTTTTCGGCCCTGCTCTCTGAGACCACTG	780
Qy	781	gacagtcacaggtgcatactctcacacgcatactccacggcgggctgtgggtctccatg	840
Db	781	GACCAGTCCACGGTGCATATCTCTCACACGCATCTCCACGGCGGCTGTGGGCTCTCCATG	840
Qy	841	atctctcggtctcaacatattcttatgacctttctgaggtttcccgggagaggttc	900
Db	841	ATCTCTTGCCCTTTCACCATATATCTTTATGCGCTTTCGAGGCTTTCCCGGGAGAGGTTTC	900
Qy	901	aagtccagaagatgccccaaagaatccacagtggccctgggtggcagcctgttctctcgaat	960
Db	901	AAGTCAGAAGATGCCCCAAAGATCCACGTGGGCCCTGGGTGGCAGCCTGTTCTCTCTGAAT	960
Qy	961	ctagcctcttggtcaatgggagtggtctcaagggtctgagtgcctgctgtggcc	1020
Db	961	CTGGCCCTTCTGTGTCAATGTGGGAGTGCTCAAGGGGCTGTGATGTGCTCTGTGGGCC	1020
Qy	1021	cgggggggctgtcttcacatacttctgtctgtgcttcacctggatgggaccttgaagcc	1080
Db	1021	CGGGGGGCTGTCTTCCACATACATTCCTGTCTGTGCTTCTACCTCGATGGGCTTGAAGCC	1080
Qy	1081	ttcacctctacctgtctgctgtcaggtgttctcaacacctacttctgggacctacttctgt	1140
Db	1081	TTCCACTCTACCTGTCTGCTGTCAAGGCTCTTCAACACTACTTCGGGCACTACTTCCCTG	1140
Qy	1141	aagctgagcctggtgggctggggcctgcgccctctgattggtcatctggcactggagtgcc	1200
Db	1141	AAGCTGAGCCTGTGTGGGCTGGGGCTGCGGCCCTGTATGGTCTATCGGCACTGGGAGTGCC	1200
Qy	1201	aacagctacggcctctacacctccgtgtagaggagaacccacactctctcggagctatgc	1260
Db	1201	AACAGCTACGGCTCTTACACCATCTGATAGGGAGAACCGACCTCTCTTGGAGCTATGC	1260
Qy	1261	tggttccgtgaagggaacacatgtacgcctcttatatcaacgttcaacgctacttctctc	1320
Db	1261	TGGTTCCGTGAAGGGACAACCATGTACCGCCTCTATATCACCGTCCACGGGCTACTTCCCTC	1320
Qy	1321	atacctctctctttggcatggtgtctctggccctgggtctctggaagatctctcaacctg	1380
Db	1321	ATCACCTTCTCTTTGGCATGGTGTCTCTGGGCCCTGGTGTCTGGAAAGATCTTTCACCCCTG	1380
Qy	1381	tccgtgctacagcgtcaagagcggggaagacaggaagaggtgctcacacctgtg	1440
Db	1381	TCCGTGTCTACACGGTCAAGAGCGGGGAAGAACCC-----GGTGTCTACCCCTGTCTG	1433
Qy	1441	ggcctctcagcctggtgggt	1461
Db	1434	GGCCTCTCGAGGCTTGCAAGT	1454

RESULT	6
AXI67305	
LOCUS	
DEFINITION	Sequence 47 from Patent WO0142287.
AXI67305	
ACCESSION	
VERSION	AXI67305.1 GI:14596747
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 1203)	
DONOHOG.G., SCOVILLE,J., TURNER,C.A., FRIEDRICH,G., ZAMBROWICZ,B. AND SANDS,A.T.	
PAT 03-JUL-2001	
linear	
DNA	
1203 bp	
AXI67305	

Db 875 TCACCTTCCTCTTTGGCATGTGGTCTCGCCCTGGTGGTCTGGAAGATCTTACCCCTGT 934
QY 1382 ccggtgctacagcgtcaagagcggggaagaaacccgaagagtgctcaacctgctgg 1441
Db 935 CCGGTGCTACACGGTCAAGGAGCGGGGGAAGAACCGGAAGGTGCTACCCCTGCTGG 994
QY 1442 gctctcagagcgtggtggtgacatggggtggtggccatcttcaacccgttggcctct 1501
Db 995 GCCTCTCGAGCCTGTGGGTGTGACATGGGGTGTGGCCATCTTACCCCGTGGCCCTCT 1054
QY 1502 ccacgtctacatctttgcaatttcaactccttgaaggggtgtcttcaactgctgctggt 1561
Db 1055 CCACCGCTTACATCTTTGACATTTTCAACTCTCTTCAAGGTGTCTATCTGCTGCTGGT 1114
QY 1562 tcacatctttacctcccaagtcaagacacacacagctcctctctactcaagattgg 1621
Db 1115 TCACCATCTTTACCTCCCAAGTCAAGACACACAGTCTCTCTCTACTTCAAGATTGG 1174
QY 1622 accaggccactccgcacatctcaagaatag 1650
Db 1175 ACCAGGCCACTCCGCATCTCAAGAAATAG 1203

RESULT 7
AX167285
LOCUS AX167285 1119 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 27 from Patent WO0142287.
ACCESSION AX167285
VERSION AX167285.1 GI:14596737
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1119)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
and Sands,A.T.
TITLE Novel human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142287-A 27 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source location/Qualifiers
1..1119
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 182 a 352 c 315 g 270 t
ORIGIN

Query Match 63.9%; Score 1054; DB 6; Length 1119;
Best Local Similarity 99.5%; Pred. No. 5.2e-207;
Matches 1057; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 482 ctctctcaaggccccggctcgccctggagatggagcggcgtgttgtaacaatgcc 541
Db 35 CCCTTTACAGGGCCCCGGCTCGGCTGGGAGATGGCAGCGCGTGTGAACAATCGCC 94
QY 542 tgggtggttgagtggtggaacaaatgcattgcacaaagctgctgaagcctctggagatcg 601
Db 95 TGTGGGTGTGAGTGTGGGCAAAATGCATGTCAACAAGTGGCTGAGCCTCTGGAGATCG 154
QY 602 tctctctcaccagcagccccccttaacatgacctcactgtgtattctgggattga 661
Db 155 TCTTCTCTCACCAGCGACCGCCCTTAACATGACCTCACCTGTGTATCTGGGATGTA 214
QY 662 ctaaaaggaccactgagagactggtctcttctgaggtgtgtccacggaggtcagacctgag 721
Db 215 CTAAGGGACCACTGGAGACTGGTCTCTCTGAGGGCTGTCTCCACGGAGGTTCAGACCTGAG 274
QY 722 ggaacctgtgctgtgacacactgaccttttttgcctgctcctgagaccacctgg 781
Db 275 GGACCGTGTGCTGCTGACCACTGACCTTTTTCGCCCTGCTCTGAGACCCACCTGG 334
QY 782 accagtcacaggttgatatactcctcacacgcacatctccacggcgtgtgggtctccatga 841

Db 335 ACCAGTCCACGGTGCATATCTCTCACACGCATCTCCAGCGGGCTGTGGGTCTCCATGA 394
QY 842 tcttcctggccttcaacattattctttatccttctgaggtttcccggaagattca 901
Db 395 TCTTCCCTGGGCTTACCATTTATCTTTATGCCCTTCTGAGGCTTTCCCGGGAGAGGTCA 454
QY 902 agtcagaagatgcccccaaatccacgtggccctgggtggcagcctgttctcctcgaatc 961
Db 455 AGTCAGAAGATGCCCAAGATCCACGTGGCCCTGGGTGGCAGCCTGTTCCTCCTGAATC 514
QY 962 tggcctcttgytgcattgtgggagtggtctcaaaagggtctgatgctgctgtggtggcc 1021
Db 515 TGGCCTTCTTGGTCAATGTGGGAGTGGCTCAAAAGGGCTCTGATGCTGCTGCTGGGCC 574
QY 1022 ggggggtgcttccactactcctgctgtgcttccactcactgagtggtggcctgaagcct 1081
Db 575 GGGGGGTGCTTCCACTACTTCTGCTGCTGCTTCACTGGATGGGCTTGAAGCCT 634
QY 1082 tccacctctactgctcgtcgtcagggcttctcaacactacttccggcactacttctcga 1141
Db 635 TCCACCTCTACCTGCTGCTGTCAGGGTCTTCAACACCTACTTCGGGSCACTACTTCTCTGA 694
QY 1142 agctgagcctggtgggtggcctgcccctctgagtggtcctgagtgagtgcca 1201
Db 695 AGCTGAGCCTGGTGGGCTGGGCGCTGCCGCCCTGTATGTCATCGGCACCTGGAGTGCCA 754
QY 1202 acagtaagcctctacacatccgtgatagggagaacccacactctctgagctatgct 1261
Db 755 ACAGTACGGCCTCTACACCATCGGTATAGGGAGAACCGACCTCTCTGGAGTATGCT 814
QY 1262 ggttcctggaaggacacacatgtacgcccctctatatcaacgctcacggctacttctca 1321
Db 815 GGTTCGTGAGGGACACACCATGTACGCCCTCTATATACCGTCCACGGCTACTTCTCTCA 874
QY 1322 tcaacttctcttggcctggtgctcctgcccctggtggtctgagaaatcttcaacctgt 1381
Db 875 TCACCTTCTCTTTGGATGTTGCTGCTGCGCCCTGGTGGTCTGGAAGATCTTCACTGCT 934
QY 1382 cccgtctacagcgtcaagagcggggaagaaacccgaagaggtgctcaccctgctg 1441
Db 935 CCCGTGCTACAGCGGTCAAGGAGCGGGGGAAGAACCCGGAAGAGTGTCTCACCCTGTGG 994
QY 1442 gcctctgagcgtggtggtgacatggggtggccatcttcaacccctgtggcctct 1501
Db 995 GCCTCTCAGCCTGTGGGTGTGACATGGGGTGTGGCCATCTTCAACCCCTGTGGGCCCTCT 1054
QY 1502 ccacgtctacatcttgcacttttcaactccttctgcaaggtg 1543
Db 1055 CCACCGTCTACATCTTTGCACTTTTCACTTTCACCTCTTTCGAAGGTG 1096

RESULT 8
AX167303
LOCUS AX167303 1020 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 45 from Patent WO0142287.
ACCESSION AX167303
VERSION AX167303.1 GI:14596746
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.
and Sands,A.T.
TITLE Novel human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142287-A 45 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source location/Qualifiers
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"

QY	1082	tcacacttacgtctgctgacaggtcttcaaacacactacttcgggaactattcctga	1141
Db	635	TCACACTTCTACCTGCTGCTGTCAGGGTCTTCAACACCTACTTCGGGCACTACTTCCTCTGA	694
QY	1142	agctgaagcctggtgggctgggctgcccgcctgatggtcatcggaactggagtgcca	1201
Db	695	AGCTGAGCCTTGGTGGGCTGGGGCTGCCCCCTGATGGTCACTCGGCACCTGGGAGTGCCA	754
QY	1202	acagctacggcctctacacacatccgtgataggagaacccgcacactctctggagctatgct	1261
Db	755	ACAGCTACGGCCTCTACACCATCCGTGATAGGAGAACCGCACCTCTCTGAGCTATGCT	814
QY	1262	ggttcctgaaggagcaacacatgtacgcccctctatatcaacgctccacggctacttcctca	1321
Db	815	GGTTCCGCTGAAGGAGCAACATGATACCCCTCTATATACCGCTCCACGGCTACTTCCTCA	874
QY	1322	tcacactctcttggcatggtgctcctggcctggcctgggtgctggaagatcttcacccctgt	1381
Db	875	TCACCTTCTCTTGGCATGGTGGTCTGGCCCTGGTGGTCTGGAAGATCTTACCCCTGT	934
QY	1382	cccgtgtacagcgggtcaaggagcgggggaagaacccggaagaggtgctcacccctgctgg	1441
Db	935	CCGCTGCTACAGCGGTCAAGAGCGGGGGAGAACCCGGAAGAGGTGCTCACCTGCTGG	994
QY	1442	gcctctgagcctggtgggt 1461	
Db	995	GCCTCTCGAGCCTTGCAAGT 1014	
RESULT 10			
LOCUS	AX167295		
DEFINITION	Sequence 37 from Patent WO0142287.		
ACCESSION	AX167295		
VERSION	AX167295.1 GI:14596742		
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1011) Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Novel human membrane proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0142287-A 37 14-JUN-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers I. .1011 /organism="Homo sapiens" /db_xref="taxon:9606"		
BASE COUNT	162 a	319 c	286 g 244 t
ORIGIN			
Query Match 57.7%; Score 952; DB 6; Length 1011;			
Best Local Similarity 98.8%; Pred. No. 5.4e-186;			
Matches 972; Conservative 0; Mismatches 5; Indels 7; Gaps 1;			
QY	482	ctctctcaaggccccggcctggcctgggagatggcagcgcggtgttggaacaatgcc	541
Db	35	CCCTTTACAGAGCCCCCGGCTGGCTGGGAGATGGCAGCGCGGTGTGAACAATCGCC	94
QY	542	tgtgtggttggatgggacaaaatgcatgtcaccaagctgctgagccttgagatcg	601
Db	95	TGTGGGTTTGTAGTGTGGGACAAATGCAATGTACCAAGCTGGCTGTGGAGATCG	154
QY	602	tcttctctaccagcagcccccttaacatgacctcactgtgtattcttggagtgga	661
Db	155	TCTTCTCTCACCAGCAGCCGCCCTTAACATGACCTCACCCTGTGTATTCGGGATGCA	214
QY	662	ctaaagggaccactgagagactggtctctgagggctgtccacgaggtcagacctgagg	721
Db	215	CTAAGGGACCACTGGAGACTGGTCTTCTGAGGGTGTCTCCACGGAGGTCAGACCTCAGG	274

QY	722	ggaccgtgtgctgtgacccaactgacctttttccctgctcctctgagacccaccttgg	781
Db	275	GGACCGTGTGCTGTGTGACCACTGACCTTTTCGCCCTGCTCTGAGACCACTGG	334
QY	782	accagtcacaggtgcatatactcaccacgcactctcccaggcgggctgtggtgtccatga	841
Db	335	ACCACTCCACGGTGCATATCTCACAGCATCTCCAGGGGGGTGTGGGTCTCCATGA	394
QY	842	tcttctggccttcacattattcttattgctttctgagccttcccgagaggttca	901
Db	395	TCCTTCCCTGGCCTTACCATTATTCTTATGCTTTTCAGGCTTTCCCGGAGAGGTCA	454
QY	902	agtccagaagatgccccaaagatccacgtggcctgggtggcagcgtgttctctcctgaac	961
Db	455	AGTCAGAAGATGCCCAAGATCCACGTGGCCCTGGGTGGCAGCCTTCTCTCTCTGAATC	514
QY	962	tggccttcttggtaaatgtgggagtgctcaaaagggtctgaatgctgctgagcct	1021
Db	515	TGGCCTTCTTGGTCAATGTGGGAGTGGCTCAAAAGGGGTCTGATGCTCCCTGTGGGCC	574
QY	1022	ggggggctgttccactacttctgctgtgcttcaactcactggtatggcctgagcct	1081
Db	575	GGGGGGCTGTCTTCCACTACTTCTGCTGTGCTTCACTGATGGGCTTGAAGCCT	634
QY	1082	tcacacttactgctgctgctcaggtgttcaaacactacttctggcactacttctga	1141
Db	635	TCCACCTTACTGCTGCTGCTCAGGGTCTTCAACACTACTTCGGGCACCTACTTCTGA	694
QY	1142	agctgagcctggtgggctggcctgcccgcctgtatggttcacgcctcagcctgagtgcca	1201
Db	695	AGCTGAGCCTGCTGGGCTGGGCTGCGGCCCTGATGTTGTCATCGGACCTGGGAGTGCCA	754
QY	1202	acagtaacggcctctaaccaatccgtgataggagagacccacacactctctgagctatgct	1261
Db	755	ACAGCTAGGGCTCTACACCATCCGTATAGGGAGAACCGCACCTCTCTGAGTATGCT	814
QY	1262	ggtccgtgaaggagacaacacatgtacgcccctctatacacgcctcagcctgacttctca	1321
Db	815	GTTTCGGGAGGAGACAAACCATGTACGCCCTCTATATACCGTCCACGGCTACTTCTCA	874
QY	1322	tcacacttcttggcatggtgctcctggcctgggtgttggatggaagatcttccacctgt	1381
Db	875	TCACCTTCTCTTGGCATGGTGGTCTGCTGGCCCTGGTGTGGAAGATCTTCAACCTGT	934
QY	1382	ccgtgtcagcgggtcaaggagcgggggaagaacccggaagaaggtgtcaccctgctgg	1441
Db	935	CCCGTGTACAGCGGTCAAGGAGCGGGGAGAAC-----GGTGTCAACCTGTCTGG	987
QY	1442	gcctctgagcctggtggtgtga 1465	
Db	988	CCCTCTCGAGCCTGTGGGTGTGA 1011	
RESULT 11			
LOCUS	AX167275	1068 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 17 from Patent WO0142287.		
ACCESSION	AX167275		
VERSION	AX167275.1 GI:14596732		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1068) Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
AUTHORS	Novel human membrane proteins and polynucleotides encoding the same		
TITLE	Patent: WO 0142287-A 17 14-JUN-2001;		
JOURNAL	Lexicon Genetics Incorporated (US)		
FEATURES	Location/Qualifiers 1. .1068		
source			

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
177 a 336 c 298 g 257 t
BASE COUNT
ORIGIN

```

Query Match	57.1%	Score 941.6;	DB 6;	Length 1068;
Best Local Similarity	98.4%;	Pred. No. 7.5e-184;		
Matches 964;	Conservative 0;	Mismatches 9;	Indels 7;	Gaps
Qy	482	ctctcttcaagggccccggctcgccctggagatggcaagcgcgtgttgaaacaatcgcc	541	
Db	35	CCCTTTCACAGGGCCCGCGCTGGCCTGGGAATGSCAGCGCGTGTGTAACAATCGCC	94	
Qy	542	tgggtgggttgagtgtaggacaatgatctcaccaaagctggctgagcctctggagatcg	601	
Db	95	TGGTGGGTTCAGTGTGGGACAAATGATCTACCAAGCTGGCTGAGCCTCTGGAGATCG	154	
Qy	602	tcttctctcaacgagcagccccctcaacatgacctcaoctgtgtattctgggatgtga	661	
Db	155	TCCTTCTCTCACACGACACGCCCCCTAACATGACCTTCACCTGTGTATCTGGGATGTGA	214	
Qy	662	ctaaaggaccactggagactggctctctgagggctgctccacgaggtccaacacctgag	721	
Db	215	CTAAGGGACCACCTGGAGACTGGTCTCTGAGGGCTGTCTCCAGGAGTTCAGACCTGAGG	274	
Qy	722	ggacggtgtgctgtgtgacacactgaccttttctgcctctgcctgagacccaacttgg	781	
Db	275	GGACCGTGTGCTGTGTGACACCTGACCTATTTTCGCCCTGTCTCTGAGACCCACCTTGG	334	
Qy	782	accagtcacggtgtcatctcctcacagcatctccacaggcggcgtgtgggtgtccatga	841	
Db	335	ACCAAGTCCACGGTGCAATCCTCACACGCACTCTCCACAGCGGGCTGTGGGGTCTCCATGA	394	
Qy	842	tcttctcggcctcaccattattcttatgtccttctgaggtcttccgggagaggttca	901	
Db	395	TCCTCCCTGGCCTTCACCATTTATCTTTATGCCTTTCTGAGGCTTTCCCGGGAGAGTTCA	454	
Qy	902	agtcagaagatgccccaaaagatccacgtggccctgggtggcagcctgttctcctcgaatc	961	
Db	455	AGTCAGAAGATGCCCCAAAAGATCCACGTGGCCCTGGGTGGCAGCCTGTTCCTTCCTGAATC	514	
Qy	962	tggcctctcttggtcaatgtgtgggagtggtccaaagggtctgatctgctgtctggggccc	1021	
Db	515	TGGCCTTCTTGGTCAATGTGGGGAGTGGCTCAAGGGGTCTGATGCTGCCTGCTGGGCC	574	
Qy	1022	ggggggcgtcttccactactctcgtctgtgtccttcacctggatggccttgaagcct	1081	
Db	575	GGGGGGCTGTCTTCCACTACTTCTCTGTGTGCCTTTCACCTGGATGGSCCTTGAAGCCT	634	
Qy	1082	tcacactctacctgcgtgtcaaggtcttcaacacctacttcgggcactacttctctga	1141	
Db	635	TCCACCTTACCTTGTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	694	
Qy	1142	agctgagcctgtggcgtggcgtggcgtggcgtggcgtggcgtggcgtggcgtggcgtggc	1201	
Db	695	AGCTGAGCCCTGGTGGCTGGCGCCCTGTGATGGTGTATCGGCCTACTGGGAGTGCCA	754	
Qy	1202	acagctacgacctctacacacatcgtgtgatgggagaaocgcaacctctctggagctatgct	1261	
Db	755	ACAGCTACGGCCCTTACACATCCGTGTATGGGAGAACCGCACCTCTCTGGAGCTATGCT	814	
Qy	1262	ggtccgtgaaggacaaacatgacgcctctatcacogtccaaggtacttctctca	1321	
Db	815	GGTTCGTGNAAGGGACACCATGTACGCCCTCTATATCACCGTCCACGGCTACTTCCCTCA	874	
Qy	1322	tcaccttccctcttggcatgggtggtcctggcctgtgtgtgtggagatcttccacctgt	1381	
Db	875	TCACCTTCCTCTTTGGCATGGTGGCTTGGCCCTTGGTGTCTGGAAGATCTTCACCTGT	934	
Qy	1382	ccggtgtacagcgttcaaggagcggggagaaacccggaagaggtgtcaccctgtcg	1441	
Db	935	CCCGTGCTACAGCGGTCAAGAGAGCGGGGAGAACCC-----GGTGCTACCTCTGTGG	987	

QY	1442	gcctcagcgccgttgggt	1461
Dd	988	GCCTCAGCGCTTGAAGT	1007
RESULT	12		
LOCUS	AXI167283	936 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 25 from Patent WO0142287.		
ACCESSION	AXI167283		
VERSION	AXI167283.1	GI:	14596736
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 936) Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B. and Sands,A.T. Novel human membrane proteins and polynucleotides encoding the same Patent: WO 0142287-A 25 14-JUN-2001;		
REFERENCE	Lexicon Genetics Incorporated (US)		
AUTHORS	Location/Qualifiers		
TITLE	1..936		
JOURNAL	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
SOURCE	153 a 289 c 263 g 231 t		
BASE COUNT			
ORIGIN			

Query Match	55.3%;	Score 913;	DB 6;	Length 936;
Best Local Similarity	100.0%;	Prod. No. 5.9e-178;		
Matches 913;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	631	atgacctcacctgtattctgggatgtgactaaaggaccactggagactggtctct	690	
Db	1	ATGACCCCTACCTGTGTATTCTGGGATGTGACTAAAGGGACCACCTGGAGACTGGTCTCT	60	
QY	691	gaggctgtctcaaggaggtcagacctgaggggaccgtgtgtctgtgacacctgacc	750	
Db	61	GAGGGCTGCTCCACGAGGTACAGACTTGAGGGGACCGTGTGCTGTGTGACCACTGACC	120	
QY	751	tttttcgcccctctctctgagacccaccttgaccagtcacaggtgcatactctcacgc	810	
Db	121	TTTTTCGCCCTCTCTCTGAGACCCACCTTGGACCACTCCACGGTGCAATCTTCACACG	180	
QY	811	atctccaggggggctgtgggtctccatgatcttcctggccttcaacattattcttat	870	
Db	181	ATCTCCCAAGCGGGCTGTGGGCTCTCCATGATCTTCTTGGCCTTCACCATATTCTTTAT	240	
QY	871	gcctttctgaggcttcccgagaggttcaagtcagaagatgccccaaagatccactg	930	
Db	241	GCCTTCTGAGGCTTTCOCGGGAGAGGTTCAAGTCAGAAGATGCCCAAGATCCACGTG	300	
QY	931	gccttgggtggagcctgttccctcgtgaatctggccttcttggtcaatgtggggagtgc	990	
Db	301	GCCTTGGGTGGCAGGCTTCTCTCTGAACTGTGGCTTCTTGGTCAATGTGGGAGTGGC	360	
QY	991	lcaaaggaggtctgactgctgtgtggccggggggtgtctctccactacttctgctc	1050	
Db	361	TCAAGGGGTCTGATGTGCTGTGTGGGCCGGGGGGGTGTCTTCCACTACTTCTGCTC	420	
QY	1051	tgtgcttccactggatgggcttgaagccttccacctctcactgctgctgtcagggtc	1110	
Db	421	TGTGCTTCACTGGATGGGCTTTGAAGCCTTCCACCTCTACCTGCTCGCTGTGAGGGTC	480	
QY	1111	ttcaacacctacttcgggcaacttctgaagctgagcctgtgtggctggggcctgccc	1170	
Db	481	TTCAACACCTACTTCGGGCACTACTTCTGAAGCTGAGCCTGTGGGTGGGGCTGCC	540	
QY	1171	gccctgatggttcacggcactgggagtgcacacagctacggcctctacacacctcgtgat	1230	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 20:20:01 ; Search time 212.41 Seconds
(without alignments)
13336.984 Million cell updates/sec

Title: US-09-733-387-43
Perfect score: 1650
Sequence: 1 atggcgacgccagggcct.....actccgcatctcaagaatag 1650

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1650	100.0	1650	22	Human novel G-prot
2	1543	93.5	1566	22	Human novel G-prot
3	1454.6	88.2	1527	22	Human novel G-prot
4	1441	87.3	1458	22	Human novel G-prot
5	1430.6	86.7	1515	22	Human novel G-prot
6	1161	70.4	1203	22	Human novel G-prot
7	1054	63.9	1119	22	Human novel G-prot
8	1020	61.8	1020	22	Human novel G-prot
9	965.6	58.5	1080	22	Human novel G-prot

10	952	57.7	1011	22	Human novel G-prot
11	941.6	57.1	1068	22	Human novel G-prot
12	913	55.3	936	22	Human novel G-prot
13	824.6	50.0	897	22	Human novel G-prot
14	811	49.2	828	22	Human novel G-prot
15	800.6	48.5	885	22	Human novel G-prot
16	772.6	46.8	825	22	Human novel G-prot
17	665.6	40.3	741	22	Human novel G-prot
18	663.4	40.2	2433	23	DNA encoding novel
19	628.4	38.1	678	22	Human novel G-prot
20	611.6	37.1	811	22	Human bone marrow
21	611.4	37.1	1755	22	Human bone marrow
22	609.4	36.9	611	22	Human EST-derived
23	577.2	35.0	702	22	Human novel G-prot
24	563.6	34.2	633	22	Human novel G-prot
25	553.2	33.5	690	22	Human novel G-prot
26	495.2	30.0	2495	21	Human secreted pro
27	495.2	30.0	2502	22	Human secreted pro
28	490	29.7	4036	22	Human novel G-prot
29	472	28.6	1431	22	Human secreted pro
30	394	23.9	612	22	Human novel G-prot
31	380.2	23.0	838	22	Human secreted pro
32	287	17.4	528	22	Human novel G-prot
33	287	17.4	584	22	Human bone marrow
34	287	17.4	4992	22	Human immune/haema
35	287	17.4	4992	22	Human immune/haema
36	284	17.2	284	22	Human bone marrow
37	266.6	16.2	603	22	Human NGPCR38 codi
38	255	15.5	536	22	Human immune/haema
39	255	15.5	536	22	Human polynucleoti
40	231.4	14.0	1193	24	Human secreted pro
41	198.6	12.0	489	22	Human novel G-prot
42	195.6	11.9	3861	22	Human immune/haema
43	195.6	11.9	3861	22	Human immune/haema
44	185	11.2	420	22	Human novel G-prot
45	174.6	10.6	477	22	Human novel G-prot

ALIGNMENTS

RESULT 1
AAD08779
ID AAD08779 standard; cdna; 1650 BP.
XX
AC AAD08779;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) #22.
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction; drug screening; transmembrane protein; mental disorder; gene therapy; nontropic; pharmacogenomics; behavioural disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS
FT 1..1650
FT /*tag= a
FT /product= "Human NGPCR protein"
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;

Sands AT;

WPI; 2001-381634/40.
P-PSDB; AAE04469.

Novel polynucleotides encoding human G protein coupled receptors useful for drug screening, diagnosis and in gene therapy of physiological or behavioral disorders.

Claim 1; Page 84; 91pp; English.

The present sequence is a cDNA encoding human novel G protein coupled receptor (NGPCR). NGPCRs are transmembrane proteins that span the cellular membrane and are involved in signal transduction after ligand binding. NGPCR DNA sequences are useful as hybridisation probes for screening libraries and assessing gene expression patterns. Sequences derived from regions adjacent to the intron/exon boundaries of NGPCR gene can be used to design primers for use in amplification assays to detect mutations within the exons, splice sites, introns that can be used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene products are useful for identifying compounds that modulate gene expression or gene product activity. Such compounds are useful in the treatment of mental, physiological or behavioural disorders and diseases. NGPCR DNA sequences and antibodies are useful for diagnostic and prognostic evaluation of disorders related to NGPCR function and for the identification of subjects having a predisposition to such disorders. NGPCR DNA sequences are also useful for drug screening and in gene therapy for modulating NGPCR expression and to produce genetically engineered host cells to express NGPCR products *in vivo*. The encoded NGPCR proteins are useful for generating antibodies, as reagents in diagnostic assays and for identifying other cellular gene products related to NGPCR.

Sequence 1650 BP; 313 A; 507 C; 447 G; 383 T; 0 other;

Query Match	100.0%;	Score 1650;	DB 22;	Length 1650;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1650;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 atggcacgcccaggggcctggggggcctgctcctgctcctcctcaggt 60

Db 1 atggcgacgccaggggccctggggggtcctcctgctcctgctccgacctcaggt 60

Qy 61 caggaaaagccaccgaaaggccaaagacacctgcctggggagcaacaatgtacgac 120

Db 61 caggaaaagccacccgaaagggccaaagaaacacacctgcctggggagcaacaacatgtacgac 120

Qy 121 atcttcaactgaatgacaaggctttgtgcttcaccaagtgcaggcagtcgggcagcgac 180

Db 121 atctcaactgaatgacaaggctttgtgcttcaccaagtcaggcagtcgggcagcgac 180

Qy 181 tcctgcaatgtggaaaacttgcagagatactggctaaactacgagggcccatctgatgaag 240

Db 181 tcctgcaatgtggaaaacttcagagatactggctaactacgaggcccatctgatgaag 240

QY 241 gaaggttgacgcagaaggagaacacgcctttcctgaaggctttggtccagaacctcagc 300

Db 241 gaaggttgacgcagaaggtaaacacgcctttcctgaaggctttggtccagaacctcagc 300

Qy 301 accaacactgcagaagacttctatttctctctgagccctctcaggttccgaggcaggtg 360

Db 301 accaactgcagaagactctatttctctctgagccctctcaggttccgaggcaggtg 360

Qy 361 atgaaggacgaggacaagccccctgacagagtgcgacttccaagagcctttttcgatcc 420

Dd 361 atgaaggacgaggacaagccccctgacagagtgcgacttccaagagcctttttcgcattcc 420

Qy 421 ctgccaggcaacaggctctgtggtccgcttgccggtcaccattctggacattggtccagg 480

Db 421 ctgccaggcaacaggctctgtggtccgccttgccgtcaccattctggacattggtccagg 480

Qy 481 actcttcaaggccccgctcgctggcctgggagatggcagcgctgttgacaatcgc 540

[illegible]

QY 1021 cggggggctgtcttcactactctctgtctgtctgtcttcactcagtgatggcctgaagcc 1080
Db 1021 cggggggctgtcttcactactctctgtctgtctgtcttcactcagtgatggcctgaagcc 1080
QY 1081 ttcactctacactgtctgtctcagggtcttcacactactctcggcactactctctg 1140
Db 1081 ttcactctacactgtctgtctcagggtcttcacactactctcggcactactctctg 1140
QY 1141 aagctgagctgtggcctgggctgcccgcctctgattggtcactggcactggagtgcc 1200
Db 1141 aagctgagctgtggcctgggctgcccgcctctgattggtcactggcactggagtgcc 1200
QY 1201 aacagctacgctctacacactcctgatagggaacacgcacactctctgagactatgc 1260
Db 1201 aacagctacgctctacacactcctgatagggaacacgcacactctctgagactatgc 1260
QY 1261 tggcttcgtgaagggaacacactgtacgcccctctatatcacgcgtccacggtactctc 1320
Db 1261 tggcttcgtgaagggaacacactgtacgcccctctatatcacgcgtccacggtactctc 1320
QY 1321 atacactctctcttgatgtgtctctgcccctctggtctggtctggaagatttcaacctg 1380
Db 1321 atacactctctcttgatgtgtctctgcccctctggtctggtctggaagatttcaacctg 1380
QY 1381 tccctgtctacagcgttcaggagcggggaagaaacccggaagaagtgtctcacctctgct 1440
Db 1381 tccctgtctacagcgttcaggagcggggaagaaacccggaagaagtgtctcacctctgct 1440
QY 1441 ggcctctcagcctgtggtgtgacatgggggttgccattctcaaccctgtggcctc 1500
Db 1441 ggcctctcagcctgtggtgtgacatgggggttgccattctcaaccctgtggcctc 1500
QY 1501 tccacgtctacatctttgcacttttcaactctcttgcaaggtg 1543
Db 1501 tccacgtctacatctttgcacttttcaactctcttgcaaggtg 1543

RESULT 3
AAD08759
ID AAD08759 standard; cDNA; 1527 BP.
XX AC AAD08759;
XX AC AAD08759;
DT 04-SEP-2001 (first entry)
XX Human novel G-protein coupled receptor (NGPCR) #2.
XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1..1527
FT CDS /*tag= a
FT /product= "Human NGPCR protein"
XX W0200142287-A2.
XX 14-JUN-2001.
XX 07-DEC-2000; 2000WO-US33241.
XX 07-DEC-1999; 99US-0169427.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI; 2001-381634/40.

DR P-PSDB; AAE04449.
XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders -
XX Claim 4; Page 62; 91pp; English.
XX The present sequence is a cDNA encoding human novel G protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX Sequence 1527 BP; 295 A; 459 C; 426 G; 347 T; 0 other;
SQ

Query Match 88.2%; Score 1454.6; DB 22; Length 1527;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1457; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 atggcagcaccagggcctggggcctgtctctgtctctctctctcctcaggt 60
Db 1 atggcagcaccagggcctggggcctgtctctgtctctctctcctcaggt 60
QY 61 caggaaaagccacccaggaaggccaaacacactcctcctgggagcaacaacatgaagc 120
Db 61 caggaaaagccacccaggaaggccaaacacactcctcctgggagcaacaacatgaagc 120
QY 121 atctcaacttgaaagcagcagcctgtgtcttcacacacagtcaggcagtcggcagcag 180
Db 121 atctcaacttgaaagcagcagcctgtgtcttcacacacagtcaggcagtcggcagcag 180
QY 181 tctcgaatgtgaaacacttgagagatactggctaaactacagagccatctgatgaag 240
Db 181 tctcgaatgtgaaacacttgagagatactggctaaactacagagccatctgatgaag 240
QY 241 gaaggtttgacgcagaggtgaacacgccttctcctgaaggtttggtccagaacctcagc 300
Db 241 gaaggtttgacgcagaggtgaacacgccttctcctgaaggtttggtccagaacctcagc 300
QY 301 accaactcgcagagcactctctctctctcctcctcctcctcctcctcctcctcctcct 360
Db 301 accaactcgcagagcactctctctctctcctcctcctcctcctcctcctcctcctcct 360
QY 361 atgaaggacgagacaagccccctgacagagtcgacttcccaagagccttttctgatcc 420
Db 361 atgaaggacgagacaagccccctgacagagtcgacttcccaagagccttttctgatcc 420
QY 421 ctgcaaggcaacaggtctgtgtcgttcgttcgttcgttcgttcgttcgttcgttcgttc 480
Db 421 ctgcaaggcaacaggtctgtgtcgttcgttcgttcgttcgttcgttcgttcgttcgttc 480
QY 481 actctctcaaggcccccgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 540
Db 481 actctctcaaggcccccgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 540

QY 541 ctggtgggttgatgtggacaaatgcatgtcaccaagctggtgagcctctgagatc 600
|||||
Db 541 ctggtgggttgatgtggacaaatgcatgtcaccaagctggtgagcctctgagatc 600
|||||
QY 601 gtcttctcacagcgagcccccctaaacatgacctcacctctgtattctggatgtg 660
|||||
Db 601 gtcttctcacagcgagcccccctaaacatgacctcacctctgtattctggatgtg 660
|||||
QY 661 actaaaggacactgagactggtctcttgaggggtgtcacagggatcagactgag 720
|||||
Db 661 actaaaggacactgagactggtctcttgaggggtgtcacagggatcagactgag 720
|||||
QY 721 gggaccgtgtgtgtgacacactgaccttttttcgacctgtctctgagacccacttg 780
|||||
Db 721 gggaccgtgtgtgtgacacactgaccttttttcgacctgtctctgagacccacttg 780
|||||
QY 781 gaccagtccacgtgcatatctctcacacgcatctcccagcggggtgtgggtctccatg 840
|||||
Db 781 gaccagtccacgtgcatatctctcacacgcatctcccagcggggtgtgggtctccatg 840
|||||
QY 841 atttctctgcttccactattattttatgctttcttctgagcgtttccgggagaggttc 900
|||||
Db 841 atttctctgcttccactattattttatgctttcttctgagcgtttccgggagaggttc 900
|||||
QY 901 aagtcagaagatgcccaagaatccacgtggccctgggtggcagacctgttctctcgaat 960
|||||
Db 901 aagtcagaagatgcccaagaatccacgtggccctgggtggcagacctgttctctcgaat 960
|||||
QY 961 ctggccttcttgatgtgggagtggtgtcctaaagggtctgagctgctcctgctgggccc 1020
|||||
Db 961 ctggccttcttgatgtgggagtggtgtcctaaagggtctgagctgctcctgctgggccc 1020
|||||
QY 1021 cggggggctgtctccactactcctgtctgtctgcttccactgagtgagccttgaagcc 1080
|||||
Db 1021 cggggggctgtctccactactcctgtctgtctgcttccactgagtgagccttgaagcc 1080
|||||
QY 1081 ttccactctacgtctgctgcaggtgtcttcaacacctactctgggcactactctctg 1140
|||||
Db 1081 ttccactctacgtctgctgcaggtgtcttcaacacctactctgggcactactctctg 1140
|||||
QY 1141 aagctgagcctgtgggtggcctgcccgcctctgagtgatggtcagctggagtgcc 1200
|||||
Db 1141 aagctgagcctgtgggtggcctgcccgcctctgagtgatggtcagctggagtgcc 1200
|||||
QY 1201 aacagctacgctctacacctccgtgataggagaacgcacacctctctgagctatgc 1260
|||||
Db 1201 aacagctacgctctacacctccgtgataggagaacgcacacctctctgagctatgc 1260
|||||
QY 1261 tgggtccgtgaaagggaacacatgtacgcccctctatatcacogtccacggctactctc 1320
|||||
Db 1261 tgggtccgtgaaagggaacacatgtacgcccctctatatcacogtccacggctactctc 1320
|||||
QY 1321 ataccctctctttggcatgtgttctgctgcccctggtggtctggaagatcttccacctg 1380
|||||
Db 1321 ataccctctctttggcatgtgttctgctgcccctggtggtctggaagatcttccacctg 1380
|||||
QY 1381 tcccgtgtcacagcgttcaaggagcggggaagaaacccggaagaaggtgtctcaccctgctg 1440
|||||
Db 1381 tcccgtgtcacagcgttcaaggagcggggaagaaacccggaagaaggtgtctcaccctgctg 1440
|||||
QY 1441 ggcctctcgagcctgtgtgggt 1461
|||||
Db 1441 ggcctctcgagcctgtgaaagt 1461
|||||

RESULT 4

AA080774
ID AA080774 standard; cDNA; 1458 BP.
XX
AC AA080774;
XX
DT 04-SEP-2001 (first entry)
XX

DE Human novel G-protein coupled receptor (NGPCR) #17.
XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1458
FT /*tag= a
FT /product= "Human NGPCR protein"
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI: 2001-381634/40.
DR P-PSDB: AAE04464.
XX
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders
XX
PS Claim 5; Page 79; 91pp; English.
XX
CC The present sequence is a cDNA encoding human novel G protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 1458 BP; 274 A; 442 C; 410 G; 332 T; 0 other;

Query Match 87.3%; Score 1441; DB 22; Length 1458;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY 1 atggcgagcagccaggggctggggccctgtctctgtctctctctccacactcaggt 60
|||||
Db 1 atggcgagcagccaggggctggggccctgtctctgtctctctctccacactcaggt 60
|||||
QY 61 caggaaaaagccacccaggaagggcccaagaacacctctcctgggagcaacaacatgtacgac 120
|||||
Db 61 caggaaaaagccacccaggaagggcccaagaacacctctcctgggagcaacaacatgtacgac 120
|||||
QY 121 atcttcaacttgaatgacaaggctttgtgttaccacaaagtgcaggcagctgggcagcgac 180
|||||

PN WO200142287-A2.
 XX 14-JUN-2001.
 XX 07-DEC-2000; 200WO-US33241.
 XX 07-DEC-1999; 99US-0169427.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX WPI; 2001-381634/40.
 XX P-PSDB; AAE04470.
 XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioral disorders -
 XX
 PS Disclosure; Page 86; 91pp; English.
 XX
 CC The present sequence is a cDNA encoding human novel G protein coupled
 CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.
 XX
 SQ Sequence 1020 BP; 172 A; 321 C; 271 G; 256 T; 0 other;

Query Match 61.8%; Score 1020; DB 22; Length 1020;
 Best Local Similarity 100.0%; Pred. No. 7e-241;
 Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 atgacctcacctgtattctggatgtgactaaaggagaccactgagactggtcttct 690
 Db 1 atgacctcacctgtattctggatgtgactaaaggagaccactgagactggtcttct 60

QY 691 gagggctgtccacggagggtcagacctgaggggaccgtgtgctgtgaccacctgacc 750
 Db 61 gagggctgtccacggagggtcagacctgaggggaccgtgtgctgtgaccacctgacc 120

QY 751 ttttgcctgtctctgagaccacctggtgaccagtcacaggtgcatctccacacgc 810
 Db 121 ttttgcctgtctctgagaccacctggtgaccagtcacaggtgcatctccacacgc 180

QY 811 atctccacggcggtgtggtgtccatgtctctcctcctcctcctcctcctccttatt 870
 Db 181 atctccacggcggtgtggtgtccatgtctctcctcctcctcctcctcctccttatt 240

QY 871 gcctttctgaggttttcccgggaggttcaagtcagaagatgcccaagatccacgtg 930
 Db 241 gcctttctgaggttttcccgggaggttcaagtcagaagatgcccaagatccacgtg 300

QY 931 gcctgggtggagcgtgtctcctcctgaatctgaccttttgggtcgaatgtggggagtg 990
 XX
 SQ

Db 301 gcctgggtggcagcgtgtctcctcctgaatcctgaccttcttggctcaatgtgggaagtgc 360
 QY 991 tcaaaggggtctgctgctgctgctggccggggggctgtcttccactacttccctgctc 1050
 Db 361 tcaaaggggtctgctgctgctgctggccggggggctgtcttccactacttccctgctc 420
 QY 1051 tgtgccttaacctggatgggcttgaagccttccaccttaaccttaacctgctgctcagggg 1110
 Db 421 tgtgccttaacctggatgggcttgaagccttccaccttaaccttaacctgctgctcagggg 480
 QY 1111 ttaaacaccttacttcggcactacttccctgaagcctgagcctgagcctggcggcctgc 1170
 Db 481 ttaaacaccttacttcggcactacttccctgaagcctgagcctggcggcctgcggcc 540
 QY 1171 gcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1230
 Db 541 gcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 600
 QY 1231 agggagaacccacctctctgctgctgctgctgctgctgctgctgctgctgctgctg 1290
 Db 601 agggagaacccacctctctgctgctgctgctgctgctgctgctgctgctgctgctg 660
 QY 1291 ctctatatcaccgtccacgctacttctcctcctcctcctcctcctcctcctcctcctc 1350
 Db 661 ctctatatcaccgtccacgctacttctcctcctcctcctcctcctcctcctcctcctc 720
 QY 1351 gcctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1410
 Db 721 gcctggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 780
 QY 1411 aagaaccgaaaggtgctcaccctgctgctgctgctgctgctgctgctgctgctgctg 1470
 Db 781 aagaaccgaaaggtgctcaccctgctgctgctgctgctgctgctgctgctgctgctg 840
 QY 1471 ggggtggcctatctcaccctgctgctgctgctgctgctgctgctgctgctgctgct 1530
 Db 841 ggggtggcctatctcaccctgctgctgctgctgctgctgctgctgctgctgctgct 900
 QY 1531 tcttgcaggggtgttctatctgctgctgctgctgctgctgctgctgctgctgctgct 1590
 Db 901 tcttgcaggggtgttctatctgctgctgctgctgctgctgctgctgctgctgctg 960
 QY 1591 accacagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1650
 Db 961 accacagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1020

RESULT 9
 AAD08761
 ID AAD08761 standard; cDNA; 1080 BP.
 XX
 AC AAD08761;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) #4.
 XX
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 KW drug screening; transmembrane protein; mental disorder; gene therapy;
 KW neurotropic; pharmacogenomics; behavioural disorder; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1080
 FT /*tag= a
 FT /product= "Human NGPCR protein"
 XX
 XX WO200142287-A2.
 XX 14-JUN-2001.
 XX 07-DEC-2000; 200WO-US33241.

XX 07-DEC-1999; 99US-0169427.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 PI WPI: 2001-381634/40.
 DR P-PSDB; AAE04451.
 XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioral disorders -
 XX Disclosure; Page 65; 91pp; English.

CC The present sequence is a cDNA encoding human novel G protein coupled
 CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.

XX Sequence 1080 BP; 183 A; 336 C; 302 G; 259 T; 0 other;
 SQ

Query Match 58.5%; Score 965.6; DB 22; Length 1080;
 Best Local Similarity 99.1%; Pred. No. 1.7e-227;
 Matches 971; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 482 ctctctcaaggcccgctgctgggagatggcagcgctgttgaaacaatcgcc 541
 DB 35 ccccttcacaggcccgctgctgggagatggcagcgctgttgaaacaatcgcc 94

QY 542 tgggtgttgatgtggacaaatcatgtcaccagctgctgagcctctgagatcg 601
 DB 95 tgggtgttgatgtgggacaaatcatgtcaccagctgctgagcctctgagatcg 154

QY 602 tctctctcaccagcgcccgctcctaaacatgacctcaccctgtatttggatgtga 661
 DB 155 tctctctcaccagcgcccgctcctaaacatgacctcaccctgtatttggatgtga 214

QY 662 ctaaaaggacactggagactgttctctgagggctgctccacggaggtcagacctgagg 721
 DB 215 ctaaaaggacactggagactgttctctgagggctgctccacggaggtcagacctgagg 274

QY 722 ggaccgtgtgctgtgacacactgaccttttccgctgctcctgagaccacacttg 781
 DB 275 ggaccgtgtgctgtgacacactgaccttttccgctgctcctgagaccacacttg 334

QY 782 accagtcacgggtgcatactcctcacacgcatactccaggcggtgtgggttcccatga 841
 DB 335 accagtcacgggtgcatactcctcacacgcatactccaggcggtgtgggttcccatga 394

QY 842 tctctctgacctcaccattatttcttctgagccttctgagccttcccgaggaggttca 901
 DB 395 tctctctgacctcaccattatttcttctgagccttctgagccttcccgaggaggttca 454

QY 902 agtcagaagatgccccaaagatccacgtggccctgggtggcagcgtgttctctcctgaatc 961
 DB 455 agtcagaagatgccccaaagatccacgtggccctgggtggcagcgtgttctctcctgaatc 514

QY 962 tggcctcttggatcaatgtgggagtgctcaaaaggggtctgtagtgcctgctgggccc 1021
 DB 515 tggcctcttggatcaatgtgggagtgctcaaaaggggtctgtagtgcctgctgggccc 574

QY 1022 ggggggctgtcttccactacttctgctgtgcttccactgagtgagccttgaagcct 1081
 DB 575 ggggggctgtcttccactacttctgctgtgcttccactgagtgagccttgaagcct 634

QY 1082 tcacactctacctgctgctgtcagggtcttcaaacacactacttggggcactacttctga 1141
 DB 635 tcacactctacctgctgctgtcagggtcttcaaacacactacttggggcactacttctga 694

QY 1142 agctgagcctggtggctggggcctgcccgcctgagtgatggctcagcactgggagtgcca 1201
 DB 695 agctgagcctggtggctggggcctgcccgcctgagtgatggctcagcactgggagtgcca 754

QY 1202 acagctacggcctctacacacatccctgtagggagaaacgcacactctctgagctatgct 1261
 DB 755 acagctacggcctctacacacatccctgtagggagaaacgcacactctctgagctatgct 814

QY 1262 ggttcgtgaaagggaacacatgtacgcccctctatatcacgcgtccacggctacttctcca 1321
 DB 815 ggttcgtgaaagggaacacatgtacgcccctctatatcacgcgtccacggctacttctcca 874

QY 1322 tcacctctcttggatggtggtgctgcccctggtggtggtggtggaagatcttcaacctgt 1381
 DB 875 tcacctctcttggatggtggtgctgcccctggtggtggtggtggaagatcttcaacctgt 934

QY 1382 cccgtgctacagcgtcaaggagcggggaagaaacggaaagagtgctcaccctgctgg 1441
 DB 935 cccgtgctacagcgtcaaggagcggggaagaaacggaaagagtgctcaccctgctgg 994

QY 1442 gcctctcagcctggtgggt 1461
 DB 995 gcctctcagcctggaagt 1014

RESULT 10
 AAD08776
 ID AAD08776 standard; cDNA; 1011 BP.
 XX AAD08776;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) #19.
 XX
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 KW drug screening; transmembrane protein; mental disorder; gene therapy;
 KW neurotropic; pharmacogenomics; behavioural disorder; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1011
 FT /*tag= a
 FT /product= "Human NGPCR protein"
 XX
 PN WO200142287-A2.
 XX
 PD 14-JUN-2001.
 XX
 XX 07-DEC-2000; 2000WO-US33241.
 XX
 PR 07-DEC-1999; 99US-0169427.
 XX
 PA (LEXI-) LEXICON GENETICS INC.

PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX
 DR WPI: 2001-381634/40.
 DR P-PSDB; AAE04466.
 XX
 PT Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioral disorders -
 XX
 PS Disclosure: Page 81-82; 91pp; English.
 XX
 CC The present sequence is a cDNA encoding human novel G protein coupled
 CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.
 XX
 SQ Sequence 1011 BP; 162 A; 319 C; 286 G; 244 T; 0 other;

Query Match 57.7%; Score 952; DB 22; Length 1011;
 Best Local Similarity 98.8%; Pred. No. 3.6e-224;
 Matches 972; Conservative 0; Mismatches 5; Indels 7; Gaps 1;

QY 482 ctctcttcaaggccccggctcgccctggagatggcagcggcgtgttgaaacaatgcc 541
 DB ||| | |||||
 QY 542 tgggtgttgatgtgggacaataatgatgtcaacagctggctgtgagctctgagatcg 601
 DB ||||| |||||
 QY 95 tgggtgttgatgtgggacaataatgatgtcaacagctggcgtgagctctgagatcg 154
 DB ||||| |||||
 QY 602 tctctctcaccagcagccccccttaacatgacctcacctgtgtattctggatgtga 661
 DB ||||| |||||
 QY 155 tctctctcaccagcagccccccttaacatgacctcacctgtgtattctggatgtga 214
 QY 662 cttaaggagaccactggagactggtcttctgagggtgctctcaccagagtcagacctgagg 721
 DB ||||| |||||
 QY 215 cttaaggagaccactggagactggtcttctgagggtgctctcaccagagtcagacctgagg 274
 QY 722 ggaacgtgtgtgtgtgaccacatgaccttttccctgctctgagaccacacttgg 781
 DB ||||| |||||
 QY 275 ggaacgtgtgtgtgtgaccacatgaccttttccctgctctgagaccacacttgg 334
 QY 782 accagtcacaggtgcatatctcaacacgcatctccacagcggagctgtgggtctccatga 841
 DB ||||| |||||
 QY 335 accagtcacaggtgcatatctcaacacgcatctccacagcggagctgtgggtctccatga 394
 QY 842 tcttctctggccttcaccattattcttattgcttcttctgaggtcttcccgagaggttca 901
 DB ||||| |||||
 QY 395 tcttctctggccttcaccattattcttattgcttcttctgaggtcttcccgagaggttca 454
 QY 902 agtcagaagatgccccaaaatcacacgtggccctgggtggcagcctgttctctctgaatc 961
 DB ||||| |||||
 QY 455 agtcagaagatgccccaaaatcacacgtggccctgggtggcagcctgttctctctgaatc 514

QY 962 tggccttcttgggtcaatgtggggagtggtcctcaaaagggtctgatgtgctgctgggccc 1021
 DB ||||| |||||
 QY 515 tggccttcttgggtcaatgtggggagtggtcctcaaaagggtctgatgtgctgctgggccc 574
 QY 1022 ggggggctgtctccactactctctctgtgcttccactgagtgaggccttgaagcct 1081
 DB ||||| |||||
 QY 575 ggggggctgtctccactactctctctgtgcttccactgagtgaggccttgaagcct 634
 QY 1082 tccacctctaccctgctgctgctcagggtcttcaacacctacttccggcactacttctctga 1141
 DB ||||| |||||
 QY 635 tccacctctaccctgctgctgctcagggtcttcaacacctacttccggcactacttctctga 694
 QY 1142 agctgagcctgttgggtgggctgcccgcctctgatgtcctcgcgactgggagtgcca 1201
 DB ||||| |||||
 QY 695 agctgagcctgttgggtgggctgcccgcctctgatgtcctcgcgactgggagtgcca 754
 QY 1202 acagctacgctctctacacatccgtgataggagaaacgcgcaactctctggagctatgct 1261
 DB ||||| |||||
 QY 755 acgctacgctctctacacatccgtgataggagaaacgcgcaactctctggagctatgct 814
 QY 1262 ggttcctgtaaggagacaacatgacgcccctctatatcacctccacggtcacttctctca 1321
 DB ||||| |||||
 QY 815 ggttcctgtaaggagacaacatgacgcccctctatatcacctccacggtcacttctctca 874
 QY 1322 tcaccttctcttggcactggtggtcctgcccctgggtgctggaagatcttccacctgt 1381
 DB ||||| |||||
 QY 875 tcaccttctcttggcactggtggtcctgcccctgggtgctggaagatcttccacctgt 934
 QY 1382 cccgtctacagcgttcaagagcggggaagaaacgcggaaggtgctcacacctgtgctg 1441
 DB ||||| |||||
 QY 935 cccgtctacagcgttcaagagcggggaagaaacgcggaaggtgctcacacctgtgctg 987
 QY 1442 gcctctcgagcctgtggtggtgga 1465
 DB ||||| |||||
 QY 988 gcctctcgagcctgtggtggtgga 1011

RESULT 11
 AAD08766
 ID AAD08766 standard; cDNA; 1068 BP.
 XX
 AC AAD08766;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) #9.
 XX
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 KW drug screening; transmembrane protein; mental disorder; gene therapy;
 KW nontropic; pharmacogenomics; behavioural disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1068
 FT /*tag= a
 FT /product= "Human NGPCR protein"
 XX
 PN WO200142287-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000WO-US33241.
 XX
 PR 07-DEC-1999; 99US-0169427.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX
 DR WPI: 2001-381634/40.
 DR P-PSDB; AAE04466.

XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioral disorders -
 XX
 PS Disclosure; Page 70; 91pp; English.
 XX
 CC The present sequence is a cDNA encoding human novel G protein coupled
 CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.
 XX
 SQ Sequence 1068 BP; 177 A; 336 C; 298 G; 257 T; 0 other;

Query Match 57.1%; Score 941.6; DB 22; Length 1068;
 Best Local Similarity 98.4%; Pred. No. 1.3e-221;
 Matches 964; Conservative 0; Mismatches 9; Indels 7; Gaps 1;

QY 482 ctctctcaaggccccggctcgccctggagatggcagcgctgttggaacaatcgcc 541
 Db 35 ccccttcacaggccccggctcgccctggagatggcagcgctgttggaacaatcgcc 94

QY 542 tgggtgggttgagtgaggacaaatgcatgtcaccaagctggctgagcctctggagatcg 601
 Db 95 tgggtgggttgagtgaggacaaatgcatgtcaccaagctggctgagcctctggagatcg 154

QY 602 tctctctacagagacgccccctaaacatgacctcaactgtattctggagatga 661
 Db 155 tctctctacagagacgccccctaaacatgacctcaactgtattctggagatga 214

QY 662 ctaaaggaccactggagactggtctctgagggctgtccacgaggtcagacactgag 721
 Db 215 ctaaaggaccactggagactggtctctgagggctgtccacgaggtcagacactgag 274

QY 722 ggaccctgtgctgtgaccacactgaccttttttgcctctgctgagaccacactgg 781
 Db 275 ggaccctgtgctgtgaccacactgaccttttttgcctctgctgagaccacactgg 334

QY 782 accagtcacaggtgcatactccacacgcatctccacgagcgctgtgggtctccatga 841
 Db 335 accagtcacaggtgcatactccacacgcatctccacgagcgctgtgggtctccatga 394

QY 842 tctctctggccttcacattattctttatgctttctgaggtcttccgggagaggttca 901
 Db 395 tctctctggccttcacattattctttatgctttctgaggtcttccgggagaggttca 454

QY 902 agtcagaagatgccccaaagatcacactggtggcctgggtggcagcctgttccctgaatc 961
 Db 455 agtcagaagatgccccaaagatcacactggtggcctgggtggcagcctgttccctgaatc 514

QY 962 tggcctctttgttcaatgtgggagtggtctcaagggtctgtatgctgctgctgggcc 1021
 Db 515 tggcctctttgttcaatgtgggagtggtctcaagggtctgtatgctgctgctgggcc 574

QY 1022 gggggggctgtcttccacttctctgtctgtgccttccactggatggggcctgaagcct 1081

Db 575 gggggggctgtcttccacttctctgtctgtcttccactggatgggcttgaagcct 634
 QY 1082 tccaccttacctgctcgctgctcaagggtcttcaaacacttctcgggcactacttctga 1141
 Db 635 tccaccttacctgctcgctgctcaagggtcttcaaacacttctcgggcactacttctga 694

QY 1142 agctgagcctggtgggctgggctgcccgcctgaggtcatcgccactggagtgcca 1201
 Db 695 agctgagcctggtgggctgggctgcccgcctgaggtcatcgccactggagtgcca 754

QY 1202 acagtacggcctctacacatccgtgataggaggagaccacacctctctggagctatgct 1261
 Db 755 acagtacggcctctacacatccgtgataggaggagaccacacctctctggagctatgct 814

QY 1262 ggttcctggaaggagacaacatgtacgcccctctctatacacgctcacgctacttctca 1321
 Db 815 ggttcctggaaggagacaacatgtacgcccctctctatacacgctcacgctacttctca 874

QY 1322 tcaccttctcttggcatggtgctgctgcccctggtggtctggaagatcttaccctgt 1381
 Db 875 tcaccttctcttggcatggtgctgctgcccctggtggtctggaagatcttaccctgt 934

QY 1382 cccgtgtctacagcgttcaagagcgggggaagaccggaagaggtgtctcaccctgtgg 1441
 Db 935 cccgtgtctacagcgttcaagagcgggggaagacc-----ggtgctcaccctgtgg 987

QY 1442 gcctctgagcctggtgggt 1461
 Db 988 gcctctgagcctggtgcaagt 1007

RESULT 12
 AAD08770
 ID AAD08770 standard; cDNA; 936 BP.
 XX
 AC AAD08770;
 DT 04-SEP-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) #13.
 XX
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 KW drug screening; transmembrane protein; mental disorder; gene therapy;
 KW nootropic; pharmacogenomics; behavioural disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..936
 FT /*tag= a
 FT /product= "Human NGPCR protein"
 XX
 WO200142287-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000WO-US33241.
 XX
 PR 07-DEC-1999; 99US-0169427.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX
 DR WPI; 2001-381634/40.
 DR P-PSDB; RAE04460.
 XX
 PT Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioral disorders -

PS Disclosure; Page 75; 91pp; English.

XX The present sequence is a cDNA encoding human novel G protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.

XX
SQ Sequence 936 BP; 153 A; 289 C; 263 G; 231 T; 0 other;

Query Match 55.3%; Score 913; DB 22; Length 936;
Best Local Similarity 100.0%; Pred. No. 1.3e-214;
Matches 913; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 631 atgacctcactgtattctggtatgtgactaaaggaccactgagactgtcttct 690
Db 1 atgacctcactgtattctggtatgtgactaaaggaccactgagactgtcttct 60
Qy 691 gagggctgtccacgaggttcagacctgaggacccgtgtgctgtgaccacctgacc 750
Db 61 gagggctgtccacgaggttcagacctgaggacccgtgtgctgtgaccacctgacc 120
Qy 751 ttttccctgtcctcctgagaccacacttgaccagtcacagtcacagtcacacgc 810
Db 121 ttttccctgtcctcctgagaccacacttgaccagtcacagtcacacgc 180
Qy 811 atctccacgagcggctgtggtctccatgtattctcgtcctcaccattattcttat 870
Db 181 atctccacgagcggctgtggtctccatgtattctcgtcctcaccattattcttat 240
Qy 871 gcccttctgagcgtttcccgaggaggttcaagtccagaagatgcccaagatccacgtg 930
Db 241 gcccttctgagcgtttcccgaggaggttcaagtccagaagatgcccaagatccacgtg 300
Qy 931 gccctgggtggcagcgtttcctcctgaatctggccttcttggatgtggggagtgcc 990
Db 301 gccctgggtggcagcgtttcctcctgaatctggccttcttggatgtggggagtgcc 360
Qy 991 tcaaaagggtctaatcgtcctcactggtggccgggggggctgtcttcacactactcctgc 1050
Db 361 tcaaaagggtctaatcgtcctcactggtggccgggggggctgtcttcacactactcctgc 420
Qy 1051 tgtgcttccacttgatggccttgaagccttccacactcactcgtcgtcaggggtc 1110
Db 421 tgtgcttccacttgatggccttgaagccttccacactcactcgtcgtcaggggtc 480
Qy 1111 tcaaacacttactcggcacttctcctgaagctgagcctggtgggctgggctgcgcc 1170
Db 481 tcaaacacttactcggcacttctcctgaagctgagcctggtgggctgggctgcgcc 540
Qy 1171 gccctgatgtcactcagcactgaggtgccaacagactcagcctctcaacactcagtgat 1230
Db 541 gccctgatgtcactcagcactgaggtgccaacagactcagcctctcaacactcagtgat 600
Qy 1231 agggagaacgcacacctctctggagctatgtcgtgttcctggaaggaggaacacctgtacgcc 1290

Db 601 agggagaacgcacacctctctggtgagctatgctgttcctggaagggaacacctgtacgcc 660
Qy 1291 ctctatatcaccgtccacagcgtactctcctcactcactcactcctctcttggcatggtggctcctg 1350
Db 661 ctctatatcaccgtccacagcgtactctcctcactcactcactcctctcttggcatggtggctcctg 720
Qy 1351 gccctggtggtctggaagatcttccacccctgctccgtgctacagcgtgcaagagcggggg 1410
Db 721 gccctggtggtctggaagatcttccacccctgctccgtgctacagcgtgcaagagcggggg 780
Qy 1411 aagaacccgaagaaggtgctcacccctgctggcctctcagcctggtgggtgtgacatgg 1470
Db 781 aagaacccgaagaaggtgctcacccctgctggcctctcagcctggtgggtgtgacatgg 840
Qy 1471 gggttgccatcttccacccctggtggcctctcagcctgctacatcttgcacttttcaac 1530
Db 841 gggttgccatcttccacccctggtggcctctcagcctgctacatcttgcacttttcaac 900
Qy 1531 tccttgcaagggtg 1543
Db 901 tccttgcaagggtg 913

RESULT 13
AAD08760
ID AAD08760 standard; cDNA; 897 BP.
XX
AC AAD08760;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) #3.
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..897
FT /tag= a
FT /product= "Human NGPCR protein"
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-381634/40.
DR P-PSDB; AAE04450.
XX
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders -
XX
PS Disclosure; Page 64; 91pp; English.
XX
CC The present sequence is a cDNA encoding human novel G protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to

CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.

XX
SQ Sequence 897 BP; 154 A; 273 C; 250 G; 220 T; 0 other;

Query Match 50.0%; Score 824.6; DB 22; Length 897;
Best Local Similarity 99.5%; Pred. No. 7, 1e-193;
Matches 827; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 631 atgacctacactgtgtattctggatgtgactaaaggaccactggagactgtctctt 690
Db 1 atgacctacactgtgtattctggatgtgactaaaggaccactggagactgtctctt 60
QY 691 gagggctgtccacgaggtcacactgagggacgtgtgctgtgacacactgacc 750
Db 61 gagggctgtccacgaggtcacactgagggacgtgtgctgtgacacactgacc 120
QY 751 tttttgcctgtcctgagaccacacttgaccagtcacagtgcatatctctcacacgc 810
Db 121 tttttgcctgtcctgagaccacacttgaccagtcacagtgcatatctctcacacgc 180
QY 811 atctccacggcgtgtggtgtccatgtatcttctggtccctccacattattctttat 870
Db 181 atctccacggcgtgtggtgtccatgtatcttctggtccctccacattattctttat 240
QY 871 gctttctgaggtcttccgggagaggttcaagtacagagatgccccaaagatccacgtg 930
Db 241 gctttctgaggtcttccgggagaggttcaagtacagagatgccccaaagatccacgtg 300
QY 931 gccctgggtggcagcctgttctctcctgaatctggtccttcttggctcaatgtggggagtgc 990
Db 301 gccctgggtggcagcctgttctctcctgaatctggtccttcttggctcaatgtggggagtgc 360
QY 991 tcaaaagggctgatgtcgtcgtcgtggtccgggggggtgtcttccactacttctgtc 1050
Db 361 tcaaaagggctgatgtcgtcgtcgtggtccgggggggtgtcttccactacttctgtc 420
QY 1051 tgtgccttcacctggatggccttgagccttcacacttcacactgctcgtcaggggtc 1110
Db 421 tgtgccttcacctggatggccttgagccttcacacttcacactgctcgtcaggggtc 480
QY 1111 ttaaacacctactctgggcactacttctgaaagctgagcctgtggtggtcgggctgcgcc 1170
Db 481 ttaaacacctactctgggcactacttctgaaagctgagcctgtggtggtcgggctgcgcc 540
QY 1171 gccctgatgtcatcgcgactggagtgccacagctacaggcctctacacacatcctgat 1230
Db 541 gccctgatgtcatcgcgactggagtgccacagctacaggcctctacacacatcctgat 600
QY 1231 agggagaacgcacactctggaagctatgctgttcctgtaaggagcaacatgtacgc 1290
Db 601 agggagaacgcacactctggaagctatgctgttcctgtaaggagcaacatgtacgc 660
QY 1291 ctctatatcccggtccacggtacttctctcatcacttctcttcttggcatgggtcctg 1350
Db 661 ctctatatcccggtccacggtacttctctcatcacttctcttcttggcatgggtcctg 720
QY 1351 gccctggtgtctggaagatcttccacctgtcccggtgtacacggtcaggagcgggg 1410
Db 721 gccctggtgtctggaagatcttccacctgtcccggtgtacacggtcaggagcgggg 780

QY 1411 aagaaccggaagaaggtgctcacctgtcggcctctcgagcctgtgggt 1461
Db 781 aagaaccggaagaaggtgctcacctgtcggcctctcgagcctgtgcaagt 831

RESULT 14

AA080775
ID AAD08775 standard; cDNA; 828 BP.
AC AAD08775;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) #18.

XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW neotropic; pharmacogenomics; behavioural disorder; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..828
FT /*tag= a
FT /product= "Human NGPCR protein"

XX WO200142287-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-US33241.

XX 07-DEC-1999; 99US-0169427.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;

XX WPI; 2001-381634/40.
DR P-PSDB; AAE04465.

XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders

XX Disclosure; Page 80; 91pp; English.

CC The present sequence is a cDNA encoding human novel G protein coupled
CC receptor (NGPCR). NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.

XX Sequence 828 BP; 133 A; 256 C; 234 G; 205 T; 0 other;

Query Match 49.2%; Score 811; DB 22; Length 828;
Best Local Similarity 99.2%; Pred. No. 1.5e-189;
Matches 828; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY 631 atgacctcaactgtgtattcttggtatgtactaaaggagaccactggagactggtcttct 690
Db 1 atgacctcaactgtgtattcttggtatgtactaaaggagaccactggagactggtcttct 60
QY 691 gagggtctctcaaggaggtcagacctgaggggacggtgtgtctgtgacctgacc 750
Db 61 gagggtctctcaaggaggtcagacctgaggggacggtgtgtctgtgacctgacc 120
QY 751 ttttcgcccgtctctcagaccaccttgaccagtcacaggtgcatactctcacacgc 810
Db 121 ttttcgcccgtctctcagaccaccttgaccagtcacaggtgcatactctcacacgc 180
QY 811 atctccaggcggcgtggtggtctccatgatctctctgacctcaccattattcttat 870
Db 181 atctccaggcggcgtggtggtctccatgatctctctgacctcaccattattcttat 240
QY 871 gcccttctgagccttcccgaggaggttcaagtcaagaatgcccaagatccacgtg 930
Db 241 gcccttctgagccttcccgaggaggttcaagtcaagaatgcccaagatccacgtg 300
QY 931 gccctggtgacagcgttctctcctgaatcgtccttcttgcattgtgggagtgcc 990
Db 301 gccctggtgacagcgttctctcctgaatcgtccttcttgcattgtgggagtgcc 360
QY 991 tcaagggtgtgtctgctgctggtggccggggggtgcttctcactacttctctc 1050
Db 361 tcaagggtgtgtctgctgctggtggccggggggtgcttctcactacttctctc 420
QY 1051 tgtgcttcaactggtggtggttgaagccttccacctctacactgtcgtgctcagggtc 1110
Db 421 tgtgcttcaactggtggtggttgaagccttccacctctacactgtcgtgctcagggtc 480
QY 1111 tcaacactactcgggcaactactcctgaagctgagcctgtggtggcgtggcctgcc 1170
Db 481 tcaacactactcgggcaactactcctgaagctgagcctgtggtggcgtggcctgcc 540
QY 1171 gccctggtgtcactggtggtggtggtggtggtggtggtggtggtggtggtggtg 1230
Db 541 gccctggtgtcactggtggtggtggtggtggtggtggtggtggtggtggtggtg 600
QY 1231 agggagaaacgacctctgagctatgctggttcgtgaaaggagaaacacctgacgcc 1290
Db 601 agggagaaacgacctctgagctatgctggttcgtgaaaggagaaacacctgacgcc 660
QY 1291 ctctataccgtccacggtcactctctcactcactcactcactcactcactcactcact 1350
Db 661 ctctataccgtccacggtcactctctcactcactcactcactcactcactcactcact 720
QY 1351 gccctggtgtcactcactcactcactcactcactcactcactcactcactcactcact 1410
Db 721 gccctggtgtcactcactcactcactcactcactcactcactcactcactcactcact 780
QY 1411 aagaacgggaaggtgtcactcactcactcactcactcactcactcactcactcactcact 1465
Db 781 aagaac-----ggtgctcaccctgctgggctctcactcactcactcactcactcact 828

RESULT 15
AAD08765
ID AAD08765 standard; cDNA; 885 BP.
XX AAD08765;
AC AAD08765;
DT 04-SEP-2001 (first entry)
XX Human novel G-protein coupled receptor (NGPCR) #8.
DE Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW

KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS I.1885
XX /*tag= a
XX /product= "Human NGPCR protein"
XX WO200142287-A2.
XX 14-JUN-2001.
XX 07-DEC-2000; 2000WO-US33241.
XX 07-DEC-1999; 99US-0169427.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
XX Sands AT;
XX WPI; 2001-381634/40.
XX P-PSDB; AAE04455.
XX Novel polynucleotides encoding human G protein coupled receptors useful
XX for drug screening, diagnosis and in gene therapy of physiological or
XX behavioral disorders.
XX Disclosure; Page 69; 91pp; English.
XX The present sequence is a cDNA encoding human novel G protein coupled
XX receptor (NGPCR). NGPCRs are transmembrane proteins that span the
XX cellular membrane and are involved in signal transduction after ligand
XX binding. NGPCR DNA sequences are useful as hybridisation probes for
XX screening libraries and assessing gene expression patterns. Sequences
XX derived from regions adjacent to the intron/exon boundaries of NGPCR
XX gene can be used to design primers for use in amplification assays to
XX detect mutations within the exons, splice sites, introns that can be
XX used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
XX products are useful for identifying compounds that modulate gene
XX expression or gene product activity. Such compounds are useful in the
XX treatment of mental, physiological or behavioural disorders and diseases.
XX NGPCR DNA sequences and antibodies are useful for diagnostic and
XX prognostic evaluation of disorders related to NGPCR function and for the
XX identification of subjects having a predisposition to such disorders.
XX NGPCR DNA sequences are also useful for drug screening and in gene
XX therapy for modulating NGPCR expression and to produce genetically
XX engineered host cells to express NGPCR products in vivo. The encoded
XX NGPCR proteins are useful for generating antibodies, as reagents in
XX diagnostic assays and for identifying other cellular gene products
XX related to NGPCR.
XX Sequence 885 BP; 148 A; 273 C; 246 G; 218 T; 0 other;

Query Match 48.5%; Score 800.6; DB 22; Length 885;
Best Local Similarity 98.7%; Pred. No. 5.6e-187;
Matches 820; Conservative 0; Mismatches 4; Indels 7; Gaps 1;

QY 631 atgacctcaactgtgtattcttggtatgtactaaaggagaccactggagactggtcttct 690
Db 1 atgacctcaactgtgtattcttggtatgtactaaaggagaccactggagactggtcttct 60
QY 691 gagggtctctcagaggtcagacctgagggacggtgtgtctgtgacctgacc 750
Db 61 gagggtctctcagaggtcagacctgagggacggtgtgtctgtgacctgacc 120
QY 751 ttttcgcccgtctctcagaccaccttgaccagtcacaggtgcatactctcacacgc 810
Db 121 ttttcgcccgtctctcagaccaccttgaccagtcacaggtgcatactctcacacgc 180

```

QY 811 atctccaggcggctgtgggtctccatgatcttctcgccctcaccattattcttat 870
Db |||||||
QY 181 atctccaggcggctgtgggtctccatgatcttctcgccctcaccattattcttat 240
Db |||||||
QY 871 gcctttctgaggtttcccgaggaggttcaagtccagaagatgccccaaagatccacgtg 930
Db |||||||
QY 241 gcctttctgaggtttcccgaggaggttcaagtccagaagatgccccaaagatccacgtg 300
Db |||||||
QY 931 gccctgggtggcagcgtcttctctcgaatctggccttcttgggtcaaatgtgggagtggc 990
Db |||||||
QY 301 gccctgggtggcagcgtcttctctcgaatctggccttcttgggtcaaatgtgggagtggc 360
Db |||||||
QY 991 tcaaaagggtctgagctgcctgctggcccgggggtgtcttccactacttctctc 1050
Db |||||||
QY 361 tcaaaagggtctgagctgcctgctggcccgggggtgtcttccactacttctctc 420
Db |||||||
QY 1051 tgtgccttcacctgagtgagcgttgaaagccttccacctctacactgctcgtcgtcagggtc 1110
Db |||||||
QY 421 tgtgccttcacctgagtgagcgttgaaagccttccacctctacactgctcgtcgtcagggtc 480
Db |||||||
QY 1111 ttcaaacactacttcgggcactactctgaagctgagcctggtggctgggcctgcc 1170
Db |||||||
QY 481 ttcaaacactacttcgggcactactctgaagctgagcctggtggctgggcctgcc 540
Db |||||||
QY 1171 gccctgagtgatcggcactgggagtgccaaacagctacggcctctacaccatccgtgat 1230
Db |||||||
QY 541 gccctgagtgatcggcactgggagtgccaaacagctacggcctctacaccatccgtgat 600
Db |||||||
QY 1231 agggagaaccgcacctctctgagctatgctggttcctggaaggagacaacccatgtacgcc 1290
Db |||||||
QY 601 agggagaaccgcacctctctgagctatgctggttcctggaaggagacaacccatgtacgcc 660
Db |||||||
QY 1291 ctctatatcacggtccagctacttctcctacactcctctcttggcatggtgctctg 1350
Db |||||||
QY 661 ctctatatcacggtccagctacttctcctacactcctctcttggcatggtgctctg 720
Db |||||||
QY 1351 gccctggtggtctggaagatcttccacctgtcccggtgtcacagcgtcaaggagcggggg 1410
Db |||||||
QY 721 gccctggtggtctggaagatcttccacctgtcccggtgtcacagcgtcaaggagcggggg 780
Db |||||||
QY 1411 aagaaccggaagaagtgctcacccctgctggccctctcagcctggtgggt 1461
Db |||||||
QY 781 aagaacc-----ggtgctcacccctgctggccctctcagcctggtgcaagt 824

```

Search completed: June 11, 2002, 21:31:16
Job time: 4275 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2002, 20:17:59 ; Search time 53.57 Seconds
(without alignments)
7565.714 Million cell updates/sec

Title: US-09-733-387-43
Perfect score: 1650
Sequence: 1 atggagagccagggcct.....actcgcattctcaagaataag 1650

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	145.8	8.8	3090	US-09-276-531-78	Sequence 78, Appl
2	85.2	5.2	669	US-09-276-531-125	Sequence 125, Appl
3	51.6	3.1	7218	US-08-232-463-14	Sequence 14, Appl
4	48.6	2.9	1931	US-09-130-114-2	Sequence 2, Appl
5	46.2	2.8	1027	US-09-276-531-24	Sequence 24, Appl
6	45.4	2.8	12588	US-08-387-942C-1	Sequence 1, Appl
7	43.6	2.6	1521	US-08-083-948-14	Sequence 14, Appl
8	43.6	2.6	1521	US-08-393-785-14	Sequence 14, Appl
9	43.6	2.6	1521	US-08-475-694-14	Sequence 14, Appl
10	43.6	2.6	1521	US-08-712-057-14	Sequence 14, Appl
11	43.6	2.6	1647	US-08-083-948-13	Sequence 13, Appl
12	43.6	2.6	1647	US-08-393-785-13	Sequence 13, Appl
13	43.6	2.6	1647	US-08-475-694-13	Sequence 13, Appl
14	43.6	2.6	1647	US-08-712-057-13	Sequence 13, Appl
15	43.6	2.6	1865	US-08-083-948-7	Sequence 7, Appl
16	43.6	2.6	1865	US-08-393-785-7	Sequence 7, Appl
17	43.6	2.6	1865	US-08-475-694-7	Sequence 7, Appl
18	43.6	2.6	1865	US-08-712-057-7	Sequence 7, Appl
19	43.4	2.6	1468	US-08-381-433A-7	Sequence 31, Appl
20	43.2	2.6	459	US-08-387-942C-31	Sequence 24, Appl
21	43.2	2.6	5894	US-08-665-259-24	Sequence 24, Appl
22	43.2	2.6	5894	US-08-762-500-24	Sequence 74, Appl
23	43.2	2.6	6525	US-08-762-500-74	Sequence 9, Appl
24	42	2.5	1288	US-08-440-856A-9	Sequence 107, App
25	41.6	2.5	3058	US-09-276-531-107	Sequence 23, Appl
26	40.8	2.5	1893	US-09-320-878-23	Sequence 3, Appl
27	40.6	2.5	975	US-09-365-150-3	Sequence 3, Appl

28	40.2	2.4	1862	1	US-07-864-475A-1	Sequence 1, Appli
29	40.2	2.4	1862	2	US-08-468-249A-1	Sequence 2, Appli
30	40.2	2.4	1863	1	US-07-864-475A-2	Sequence 2, Appli
31	40.2	2.4	1863	2	US-08-468-249A-2	Sequence 2, Appli
32	40.2	2.4	3271	2	US-08-852-806-1	Sequence 1, Appli
33	40.2	2.4	3271	3	US-09-163-689-1	Sequence 1, Appli
34	39.8	2.4	756	3	US-08-513-974B-53	Sequence 53, Appli
35	39.8	2.4	810	3	US-08-513-974B-359	Sequence 359, App
36	39.8	2.4	1161	1	US-08-086-439C-2	Sequence 2, Appli
37	39.8	2.4	1161	1	US-08-434-877-2	Sequence 2, Appli
38	39.8	2.4	1367	3	US-08-475-742-3	Sequence 3, Appli
39	39.6	2.4	390	4	US-09-197-649-7	Sequence 7, Appli
40	39.2	2.4	5698	1	US-08-761-258-11	Sequence 11, Appl
41	39.2	2.4	5698	2	US-08-977-308-11	Sequence 11, Appl
42	38.6	2.3	2150	2	US-08-861-464-13	Sequence 13, Appl
43	38.6	2.3	2150	2	US-08-396-001-13	Sequence 13, Appl
44	38.6	2.3	2150	4	US-09-323-433A-13	Sequence 13, Appl
45	38.6	2.3	2207	4	US-08-956-322-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-276-531-78
; Sequence 78, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 853-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 3090 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14

CLONE: 1595762
US-09-276-531-78

Query Match 8.8%; Score 145.8; DB 4; Length 3090;
Best Local Similarity 54.1%; Pred. No. 2.3e-27;
Matches 372; Conservative 0; Mismatches 297; Indels 18; Gaps 3;

```
Qy 515 atgacagcggcgtgttgaacaatcgcctggtggttggatgtggaacaaatgcattgca 574
  || ||||| || || || || || || || || || || || || || || || || || ||
Db 438 ATTCCAGCAAGCTCTGGGTGAGAAGGCTCTTGGGGATTGTGGTACAGAACCAAGTAG 497
  || || || || || || || || || || || || || || || || || || || || ||
Qy 575 ccaagctggctgagcctctgagatgctctctcaccagcagcgcgccttaacatga 634
  || || || || || || || || || || || || || || || || || || || || ||
Db 498 CCAACCTCAGGGAGCGTGCTGCTCACCTTCAGGACACAGCTACAGCGAAGATGGA 557
  || || || || || || || || || || || || || || || || || || || || ||
Qy 635 ccctcaacctgtatttggatgtg-----actaaaggacacctggagactggttt 688
  || || || || || || || || || || || || || || || || || || || || ||
Db 558 CTCTGCAATGTGTTCTGGGTTCAAGACCCACATTGAGCAGCCCGGGCATTTGGAGCA 617
  || || || || || || || || || || || || || || || || || || || || ||
Qy 689 ctgagggctgctccagcagaggtcagacctgaggggacgtgctgctgtgacacctga 748
  || || || || || || || || || || || || || || || || || || || || ||
Db 618 GTGCTGGGTGAGACCGTCAGAGAGAAACCAACATCCTGCTCTGCAACCACTTGA 677
  || || || || || || || || || || || || || || || || || || || || ||
Qy 749 ccttttcgcctgctcctgagaccacacttggaccagtcacagtgatata-----tcc 802
  || || || || || || || || || || || || || || || || || || || || ||
Db 678 CCACTTTGCAAGTGTGATGCTCTCCTGGTGGAGGTGACGCGCTGCAACAGCACTACC 737
  || || || || || || || || || || || || || || || || || || || || ||
Qy 803 tcacacgcatctccagcggcgtggtgggtctccatgctctcctggccttcaccatta 862
  || || || || || || || || || || || || || || || || || || || || ||
Db 738 TGAGCCCTCTCTAGCTGGGTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
  || || || || || || || || || || || || || || || || || || || || ||
Qy 863 ttcttatgcttctgaggtcttcgcgggagaggttcaagtca-----gaagatgcc 916
  || || || || || || || || || || || || || || || || || || || || ||
Db 798 CCGCTACCTCTGCTCCAGGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
  || || || || || || || || || || || || || || || || || || || || ||
Qy 917 caaagatccacgtggcctggtggtgagcctgtctctctgaatctggtcttctgttca 976
  || || || || || || || || || || || || || || || || || || || || ||
Db 858 TCAAGGTGCACATGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
  || || || || || || || || || || || || || || || || || || || || ||
Qy 977 atgtgggagtggtcctaaagggtgtgctgctgctgctgctgctgctgctgctgcttcc 1036
  || || || || || || || || || || || || || || || || || || || || ||
Db 918 GCGAGCGGGTGGCCCTGACAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGCTG 977
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1037 actactctgctgctgctgctcctcactgagccttgaagccttcaacctctacctgc 1096
  || || || || || || || || || || || || || || || || || || || || ||
Db 978 ACTTCTCCCTGCTCACCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1097 tcgctgtcaggggttcaacacactacttcgggcacttctcctgaagctgagcctgtgtg 1156
  || || || || || || || || || || || || || || || || || || || || ||
Db 1038 TCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCATGG 1097
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1157 gctggggcctgcccgcctgattgca 1183
  || || || || || || || || || || || || || || || || || || || || ||
Db 1098 GCTGGGCTTCCCATCTTTCTGGTGA 1124
  || || || || || || || || || || || || || || || || || || || || ||
```

RESULT 2

US-09-276-531-125
; Sequence 125, Application US/09276531
; Patent No. 618398
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murty, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THYRN09
CLONE: 2715440
US-09-276-531-125

Query Match 5.2%; Score 85.2; DB 4; Length 669;
Best Local Similarity 55.9%; Pred. No. 2.2e-12;
Matches 162; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```
Qy 892 gagagggtcaagtcagaagatgcccaagaatccacgtggccctgggtggcagcctgttc 951
  || || || || || || || || || || || || || || || || || || || || ||
Db 76 GAGAAATTCGAAGGAGGATTATCCCTCCAAATCTTGATGAACCTGAGCACAGCCCTGTG 135
  || || || || || || || || || || || || || || || || || || || || ||
Qy 952 ctoctgaatctggcctctcttggtaaatgtygggagtyggcctcaaaagggtctgagctgcc 1011
  || || || || || || || || || || || || || || || || || || || || ||
Db 136 TTCTGGAATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 195
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1012 tgcctggcggcgggggctgctctccactactctcctgctgctgctcaccctgagtgaggc 1071
  || || || || || || || || || || || || || || || || || || || || ||
Db 196 TGCATTGCTGTGCAGTCTCTGTGCAATTTCTTCTCTCTGCGCAACCTTTTACCTGGATGGG 255
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1072 cttgaagccttcacacctctacacctgctgctgctgaggggtttcaacacctactctggggcac 1131
  || || || || || || || || || || || || || || || || || || || || ||
Db 256 CTAGAAGCAATTCACATGTACATTGCTCTAGTTAAAGTATTTAACACTTACACTTCGCGGA 315
  || || || || || || || || || || || || || || || || || || || || ||
Qy 1132 tactctcaagctgagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 1181
  || || || || || || || || || || || || || || || || || || || || ||
Db 316 TACATTCTAAATTCGATCAITGGCTGGGGTTTGGCTGCTTGTAGTGGT 365
  || || || || || || || || || || || || || || || || || || || || ||
```

RESULT 3

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52

[illegible]

RESULT 8
US-08-393-785-14
; Sequence 14, Application US/08393785
; Patent No. 5554369
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.
; APPLICANT: Jennings, Michael G.
; APPLICANT: Purcell, John P.
; APPLICANT: Sammons, Robert D.
; TITLE OF INVENTION: Method of Controlling Insects
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsant Company, B84F
; STREET: 700 Chesterfield Parkway No. 5554369th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,785
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/083,948
; FILING DATE: 28-JUN-1993
; APPLICATION NUMBER: US 07/762682
; FILING DATE: 23-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937195
; FILING DATE: 09-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.

REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-393-785-14

Query Match 2.6%; Score 43.6; DB 1; Length 1521;
Best Local Similarity 47.2%; Pred. No. 0.086;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1109 tttcaaacactacttggcgcactacttctgaagctgagcctggtggcgtgggctgc 1168
Db 845 TCTTCTCGCGCGCGGAGCTCGCTGCAATTGAATGCTGCTGCGCGCGCGGAGACCG 904
QY 1169 cgcgcctgatgttcacgtgagtgagtgccacagctacgctctcacaccatcgtg 1228
Db 905 GCACCTGCGCGCGCTCAGCTCCGAGATCGCGCGCGCTGGGCGCGCCCAACGCAACATCA 964
QY 1229 ataggggaaacccgcactctctgtgagctatgtgttcctggaagggaacacatgacg 1288
Db 965 TGACCGCGCGCGGACCATGTGTGAACCCACGCGGCGAGCAGTGTGATCCCG 1024
QY 1289 cctctatataccgtccacgctacttctctcctcctcctctcttctgtgcatgtgttc 1348
Db 1025 CCCTCGGATCGAGCTGGGAGAACCCCGACACCCCGCTTCGCGGAGATAGCCCCCA 1084
QY 1349 tggcctgtgtgtggaagatcttcacccctgtcccgctgcta 1390
Db 1085 TGCGCGCGCGCTCGAGACCTGGGTGAGCTTACCTGGCCA 1126

RESULT 9
US-08-475-694-14
; Sequence 14, Application US/08475694
; Patent No. 5558862
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.
; APPLICANT: Jennings, Michael G.
; APPLICANT: Purcell, John P.
; APPLICANT: Sammons, Robert D.
; TITLE OF INVENTION: Method of Controlling Insects
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5558862th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,694
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,785
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US 08/083,948
; FILING DATE: 28-JUN-1993

APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-475-694-14

Query Match 2.6%; Score 43.6; DB 1; Length 1521;
Best Local Similarity 47.2%; Pred. No. 0.086;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1109 tttcaaacactacttggcgcactacttctgaagctgagcctggtggcgtgggctgc 1168
Db 845 TCTTCTCGCGCGCGGAGCTCGCTGCAATTGAATGCTGCTGCGCGCGCGGAGACCG 904
QY 1169 cgcgcctgatgttcacgtgagtgagtgccacagctacgctctcacaccatcgtg 1228
Db 905 GCACCTGCGCGCGCTCAGCTCCGAGATCGCGCGCGCTGGGCGCGCCCAACGCAACATCA 964
QY 1229 ataggggaaacccgcactctctgtgagctatgtgttcctggaagggaacacatgacg 1288
Db 965 TGACCGCGCGCGGACCATGTGTGAACCCACGCGGCGAGCAGTGTGATCCCG 1024
QY 1289 cctctatataccgtccacgctacttctctcctcctcctcttctgtgcatgtgttc 1348
Db 1025 CCCTCGGATCGAGCTGGGAGAACCCCGACACCCCGCTTCGCGGAGATAGCCCCCA 1084
QY 1349 tggcctgtgtgtggaagatcttcacccctgtcccgctgcta 1390
Db 1085 TGCGCGCGCGCTCGAGACCTGGGTGAGCTTACCTGGCCA 1126

RESULT 10
US-08-712-057-14
; Sequence 14, Application US/08712057
; Patent No. 5763245
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.
; APPLICANT: Purcell, John P.
; APPLICANT: Pershing, Jay
; TITLE OF INVENTION: Method of Controlling Insects
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5763245th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,057
; FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-712-057-14

Query Match 2.6%; Score 43.6; DB 1; Length 1521;
Best Local Similarity 47.2%; Pred. No. 0.086;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1109 tcttaaacacctacttcggcacttctgaagctgagcctggtgggtgggcctgc 1168
Db 845 TCTTCTCGGCGCGGCGAGCTCGGCTCCATTGAAGTGTGCTGGCGCGCGGAGACGG 904
QY 1169 ccgcccctggtgctacgcgcaactggtggagtgccaaacagctacgctctacacaccctgc 1228
Db 905 GCACCTGCGCGCGCTCAGCTCGAGATCGGCGCGGCTGGGGCCCCCAACGGCAATCA 964
QY 1229 ataggagaaacgcacacctctgtgagctatgtgttcctggaagggaacacacatgtacg 1288
Db 965 TGACCGCGCGCGCGCAACCATGTGTGGAACCGGCGAGCAAGCAGTGTGATCCCGG 1024
QY 1289 ccctctatatcacggtccacgctacttctctcacccttcttcttcttcttcttcttcttct 1348
Db 1025 CCCTCGGATCGAGACTGGGACAAACCGGACAAACCGGCTCTCGCGGAGATAGCCCCCA 1084
QY 1349 tggcccctggtggttgaagatcttccacctgtcccggtgcta 1390
Db 1085 TGCGCGCGGCGCTCGAGACCTGGGTGAGCTTACCTGAGCA 1126

RESULT 11
US-08-712-057-14
Sequence 13, Application US/08083948
Patent No. 5518908
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Greenplate, John T.
APPLICANT: Jennings, Michael G.
APPLICANT: Purcell, John P.
APPLICANT: Sammons, Robert D.
TITLE OF INVENTION: Method of Controlling Insects
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
STREET: 700 Chesterfield Parkway No. 5518908th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,948
FILING DATE: 19930628
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1647 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-083-948-13

Query Match 2.6%; Score 43.6; DB 1; Length 1647;
Best Local Similarity 47.2%; Pred. No. 0.089;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 1109 tcttaaacacctacttcggcacttctgaagctgagcctggtgggtgggcctgc 1168
Db 971 TCTTCTCGGCGCGGCGAGCTCGGCTCCATTGAAGTGTGCTGGCGCGCGGAGACGG 1030
QY 1169 ccgcccctggtgctacgcgcaactggtggagtgccaaacagctacgctctacacaccctgc 1228
Db 1031 GCACCTGCGCGCGGCTCAGCTCGAGATCGGCGCGGCTGGGGCCCCCAACGGCAATCA 1090
QY 1229 ataggagaaacgcacacctctgtgagctatgtgttcctggaagggaacacacatgtacg 1288
Db 1091 TGACCGCGCGCGCGCAACCATGTGTGGAACCGGCGAGCAAGCAGTGTGATCCCGG 1150
QY 1289 ccctctatatcacggtccacgctacttctctcacccttcttcttcttcttcttcttcttct 1348
Db 1151 CCCTCGGATCGAGACTGGGACAAACCGGACAAACCGGCTCTCGCGGAGATAGCCCCCA 1210
QY 1349 tggcccctggtggttgaagatcttccacctgtcccggtgcta 1390
Db 1211 TGCGCGCGGCGCTCGAGACCTGGGTGAGCTTACCTGAGCA 1252

RESULT 12
US-08-393-785-13
Sequence 13, Application US/08393785
Patent No. 5554369
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Greenplate, John T.
APPLICANT: Jennings, Michael G.
APPLICANT: Purcell, John P.
APPLICANT: Sammons, Robert D.
TITLE OF INVENTION: Method of Controlling Insects
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
STREET: 700 Chesterfield Parkway No. 5554369th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,785
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/083,948
FILING DATE: 28-JUN-1993
APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1647 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-793-785-13

Query Match 2.6%; Score 43.6; DB 1; Length 1647;
Best Local Similarity 47.2%; Pred. No. 0.089;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 1109 tcttcaacacactcttgggacacttctgaagctgagcctggggtgggctgc 1168
DB 971 TCTTCTCGCGCGGCGAGCTCGGCTCCATTGAAGTCTGCTGCGCGCGGAGACCG 1030
QY 1169 cgcgcctgatggtatcgccactgagtgagtgccacagctacagcctacacccatgc 1228
DB 1031 GCACCCCTGCCCGGCTCAGCTCCGAGATCGGGGGGCTGGGGCCCAACGCAACATCA 1090
QY 1229 atagggagaacccacactctctgagctatgctggttcctggaaggaacacatgtac 1288
DB 1091 TGACCGCGCGGCGCAACCATGTGTGGAAACCCACGCGGCGAGCAAGCAAGTCTGCTGATCCCCG 1150
QY 1289 ccctctataacgctccacggtacttctctatcacttctcttcttggcatggtgtcc 1348
DB 1151 CCTCGGCATCGAGACTGGGACAAACCCGACAAACCCGCTTTCGCGGAGATAGCCCCCA 1210
QY 1349 tggcctgtggttctggaagatcttaccctgtcccggtgcta 1390
DB 1211 TGCGGCGGCGCTCGAGACCTGGGTGAGCTCTACCTGGCCA 1252

RESULT 13
US-08-733-387-43
Sequence 13, Application US/08/75694
Patent No. 5558862
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Greenplate, John T.
APPLICANT: Jennings, Michael G.
APPLICANT: Purcell, John P.
APPLICANT: Sammons, Robert D.
TITLE OF INVENTION: Method of Controlling Insects
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
STREET: 700 Chesterfield Parkway No. 5558862th
CITY: St. Louis

STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,694
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,785
FILING DATE: 24-FEB-1995
APPLICATION NUMBER: US 08/083,948
FILING DATE: 28-JUN-1993
APPLICATION NUMBER: US 07/762682
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937195
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10631)A
TELEPHONE: (314)537-7286
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1647 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-475-694-13

Query Match 2.6%; Score 43.6; DB 1; Length 1647;
Best Local Similarity 47.2%; Pred. No. 0.089;
Matches 133; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 1109 tcttcaacacactcttgggacacttctgaagctgagcctggggtgggctgc 1168
DB 971 TCTTCTCGCGCGGCGAGCTCGGCTCCATTGAAGTCTGCTGCGCGCGGAGACCG 1030
QY 1169 cgcgcctgatggtatcgccactgagtgagtgccacagctacagcctacacccatgc 1228
DB 1031 GCACCCCTGCCCGGCTCAGCTCCGAGATCGGGGGGCTGGGGCCCAACGCAACATCA 1090
QY 1229 atagggagaacccacactctctgagctatgctggttcctggaaggaacacatgtac 1288
DB 1091 TGACCGCGCGGCGCAACCATGTGTGGAAACCCACGCGGCGAGCAAGCAAGTCTGCTGATCCCCG 1150
QY 1289 ccctctataacgctccacggtacttctctatcacttctcttcttggcatggtgtcc 1348
DB 1151 CCTCGGCATCGAGACTGGGACAAACCCGACAAACCCGCTTTCGCGGAGATAGCCCCCA 1210
QY 1349 tggcctgtggttctggaagatcttaccctgtcccggtgcta 1390
DB 1211 TGCGGCGGCGCTCGAGACCTGGGTGAGCTCTACCTGGCCA 1252

RESULT 14
US-08-712-057-13
Sequence 13, Application US/08/12057
Patent No. 5763245
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Greenplate, John T.
APPLICANT: Purcell, John P.
APPLICANT: Pershing, Jay

RESULT 15
US-08-083-948-7
; Sequence 7, Application US/08083948
; Patent No. 5518908
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Greenplate, John T.

Search completed: June 11, 2002, 21:27:35
Job time: 4176 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 21:26:13 ; Search time 61.01 seconds
(without alignments)
999.500 Million cell updates/sec

Title: US-09-733-387-44

Perfect score: 2878

Sequence: 1 MATPRGLGALLLLLLLPTSG.....STTVSSSTARLDQASASQE 549

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2878	100.0	549	22 AAE04469	Human novel G-prot
2	2696	93.7	521	22 AAE04459	Human novel G-prot
3	2548	88.5	508	22 AAE04449	Human novel G-prot
4	2494	86.7	485	22 AAE04464	Human novel G-prot
5	2494	86.7	504	22 AAE04454	Human novel G-prot
6	2022	70.3	400	22 AAE04471	Human novel G-prot
7	1840	63.9	372	22 AAE04461	Human novel G-prot
8	1778	61.8	339	22 AAE04470	Human novel G-prot
9	1692	58.8	359	22 AAE04451	Human novel G-prot
10	1638	56.9	336	22 AAE04466	Human novel G-prot
11	1638	56.9	355	22 AAE04456	Human novel G-prot

12	1596	55.5	311	22 AAE04460	Human novel G-prot
13	1448	50.3	298	22 AAE04450	Human novel G-prot
14	1394	48.4	275	22 AAE04465	Human novel G-prot
15	1394	48.4	294	22 AAE04455	Human novel G-prot
16	1330	46.2	274	22 AAE04472	Human novel G-prot
17	1148	39.9	246	22 AAE04462	Human novel G-prot
18	1146	39.8	630	22 ABG14818	Novel human diagno
19	1094	38.0	225	22 AAE04448	Human novel G-prot
20	1089.5	37.9	585	22 ABH12463	Human bone marrow
21	1000	34.7	233	22 AAE04452	Human novel G-prot
22	946	32.9	210	22 AAE04467	Human novel G-prot
23	946	32.9	229	22 AAE04457	Human novel G-prot
24	896	31.1	211	22 ABB12380	Human bone marrow
25	804	27.9	157	22 AAE03245	Human gene 3 encod
26	799	27.8	158	22 AAE03244	Human gene 3 encod
27	671.5	23.3	203	22 AAE04473	Human novel G-prot
28	643.5	22.4	687	21 AAY79182	Haematopoietic ste
29	617	21.4	122	22 AAE03204	Human gene 3 encod
30	610.5	21.2	1013	21 AAB01247	Human HE6 receptor
31	610.5	21.2	1038	18 AAW36903	Human epididymis-s
32	607	21.1	128	22 AAE03246	Human gene 3 encod
33	603.5	21.0	512	22 AAE03246	Human BBSR seven t
34	603.5	21.0	687	21 AAB21700	Human TM7XN1. Hom
35	602.5	20.9	693	21 AAB01246	Human seven transm
36	599.5	20.8	687	20 AAY02382	Human polypeptide
37	599.5	20.8	687	22 AAM93662	Human polypeptide
38	596.5	20.7	693	20 AAY41765	Human PRO1083 prot
39	596.5	20.7	693	21 AAB44321	Human PRO1083 (UNQ
40	596.5	20.7	693	21 AAB33435	Human PRO1083 prot
41	596.5	20.7	693	21 AAY92362	G protein-coupled
42	596.5	20.7	693	22 AAM78369	Human protein SEQ
43	596.5	20.7	693	22 AAU12374	Human PRO1083 poly
44	594.5	20.7	693	20 AAY05730	Human G protein co
45	591.5	20.6	687	22 AAB88463	Human membrane or

ALIGNMENTS

RESULT 1

AAE04469
ID AAE04469 standard; Protein; 549 AA.

XX AAE04469;

XX AC
DT 04-SEP-2001 (first entry)

XX DE Human novel G-protein coupled receptor (NGPCR) protein #22.

XX KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
XX KW drug screening; transmembrane protein; mental disorder; gene therapy;
XX KW nontropic; pharmacogenomics; behavioural disorder.

XX OS Homo sapiens.

XX PN WO200142287-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-US33241.

XX PR 07-DEC-1999; 99US-0169427.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;

XX PT Sands AT;

XX DR WPI: 2001-381634/40.

XX N-PSDB; AAD08779.

XX PT Novel polynucleotides encoding human G protein coupled receptors useful
for drug screening, diagnosis and in gene therapy of physiological or

PT behavioral disorders -

PS Claim 2; Page 84-86; 91pp; English.

XX The present sequence is human novel G protein coupled receptor
XX (NGPCR) protein. NGPCRs are transmembrane proteins that span the
XX cellular membrane and are involved in signal transduction after ligand
XX binding. NGPCR DNA sequences are useful as hybridisation probes for
XX screening libraries and assessing gene expression patterns. Sequences
XX derived from regions adjacent to the intron/exon boundaries of NGPCR
XX gene can be used to design primers for use in amplification assays to
XX detect mutations within the exons, splice sites, introns that can be
XX used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
XX products are useful for identifying compounds that modulate gene
XX expression or gene product activity. Such compounds are useful in the
XX treatment of mental, physiological or behavioural disorders and diseases.
XX NGPCR DNA sequences and antibodies are useful for diagnostic and
XX prognostic evaluation of disorders related to NGPCR function and for the
XX identification of subjects having a predisposition to such disorders.
XX NGPCR DNA sequences are also useful for drug screening and in gene
XX therapy for modulating NGPCR expression and to produce genetically
XX engineered host cells to express NGPCR products in vivo. The encoded
XX NGPCR proteins are useful for generating antibodies, as reagents in
XX diagnostic assays and for identifying other cellular gene products
XX related to NGPCR.

SQ Sequence 549 AA;

Query Match 100.0%; Score 2878; DB 22; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.6e-289;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATPRGLGALLLLPTSGQEKTEGPRNTCLGSNNMYDIFNLNDKALCFTKCRQSGSD 60
Db 1 matprglgalllllptsgqektegrntclgsnnmydifnlndkalcftkcrqsgsd 60
Qy 61 SCNVENLQRYWLNVEAHLMEKGLTQKVNTFFLKALVONLSTNTAEDFYFSLPSQVPRQV 120
Db 61 scnvenlqrywlnveahlmekgltkvntfflkalvnlstntaedfyfslpsqvprqv 120
Qy 121 MKDEKPPDRVRLPKSLRSLPGNRSVVRVLAFTLIDIGPGLGPRGLGDSGVNLNR 180
Db 121 mkdekppdrvrlpkslrslpgnrsvvravlftldigpglgrglgdsgvnlmr 180
Qy 181 LVLSVGOMHVTKLAEPLEIVFHSQRPPNMTLTCVFMWVTGKTGDSWSEGCSTEVRPE 240
Db 181 lvlsvgomhvtklaepileivfshqrppnmtltcvfwmvvtgktgdswsegcstevrpe 240
Y 241 GTVCCDHLTFALLRPTLDQSTVHLITRISQAGCVSMIFLAFTILLYAFILRSERF 300
Db 241 gtvccdhltfllrptldqstvhlttrisqagcvsmiflaftillyafilrserf 300
Qy 301 KSEDAPKIHVALGSLFLNLAFLVNVGSGKSDAACWARGAVFHVFLLCFTWMLGA 360
Db 301 ksedapkihalgsflnlaflvnvsgskgsdaacwargavfhlcaftwmlga 360
Qy 361 FHLVLLAVRVNTVFHYFKLSLVGWLCPALMVIGTGSANSYGLXITRDRENTSLELC 420
Db 361 fhllvllavrvntvfhyfklslvwgwlpalmvigtgsansyglytirdrentslelc 420
Qy 421 WFRGTTMYALYIVHGYFLITFLFGWVLAFLVVKLFTLSRATAVKERNKKVLTLL 480
Db 421 wfrgttmyalyivhgyflitflfgwvllalvkvkltlsratavkernkkvltll 480
Qy 481 GLSSLVGTWGLAFTPLGLSTVYIFALENSLOGVFICWFTILYLPQSQTTSSTARL 540
Db 481 glsslvgtwglaftplglstvyifalfnslqgvficcwftillylpqsqttsstarl 540
Qy 541 DQAHSAQOE 549
Db 541 dqahsaqoe 549

RESULT 2

AAE04459
ID AAE04459 standard; Protein; 521 AA.XX
AC AAE04459;XX
DT 04-SEP-2001 (first entry)XX
DE Human novel G-protein coupled receptor (NGPCR) protein #12.

XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.

XX
OS Homo sapiens.XX
PN WO200142287-A2.XX
PD 14-JUN-2001.XX
PF 07-DEC-2000; 2000WO-US33241.XX
PR 07-DEC-1999; 99US-0169427.XX
PA (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;

XX
WPI; 2001-381634/40.DR
N-PSDB; AAD08769.

XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioural disorders -

XX
PS Disclosure: Page 73-74; 91pp; English.

XX The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.

XX
SQ Sequence 521 AA;

Query Match 93.7%; Score 2696; DB 22; Length 521;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATPRGLGALLLLPTSGQEKTEGPRNTCLGSNNMYDIFNLNDKALCFTKCRQSGSD 60

Db 1 matprglgalllllptsgqektegrntclgsnnmydifnlndkalcftkcrqsgsd 60

Qy 61 SCNVENLQRYWLNVEAHLMEKGLTQKVNTFFLKALVONLSTNTAEDFYFSLPSQVPRQV 120

||||| scvenlqrywlyneahlmkegltqkntpfkalkvnlstaeafyfslepsqvprqv 120
Db
QY 121 MKDEKPPDRVRLPKSLFRSLPCNRSVVRVLAVTILDIGPGTLFGKPRGLGDSGVLLNR 180
Db
121 mkdedkppdrvr.lpkslfrslpcnrsvvr.lavtildigpgtl.fkprglgdsgvllnr 180
QY 181 LVGLSVGQHVHTKLAEPLEIVFSHQRPNNMTLTCVFWDVTKGTTGDMSSGCSFEVRPE 240
Db
181 lvglsvghvhtklaepleivfshqrpnnmtlctcvfwdvktgtdmssescstevrpe 240
QY 241 GTVCCDHLTFFALLRPTLDQSTVHLTRISQAGCGVSMIFLAFTIILYAFRLSRERF 300
Db
241 gtvcccdhlthfllrptldgstvhltrisoagcgvsmlfraftiilyafrlsrerf 300
QY 301 KSEDAPKIHVALGGSFLNLNLAFLVNVGSGSGSDAACWARGAVHFYLLCAFTWMGLEA 360
Db
301 ksedapkihnvalggsflnlalnlaflvnnvsgsgsdacwargavhfylcaftwmglea 360
QY 361 FHLYLLAVRVNTYFGHYFLKLSLVGGLPALMVIGTGSANSYGLYTIIRDRENTSLELC 420
Db
361 fhlyllavrvntyfghyflklslvgglpalmvigtgsansyglytirdrenrtslelc 420
QY 421 WFRGTTMYALYITVHGYSFLITFLFGMVVVALVVMVKIFTLRSATAVKRGKRRKVVLTLL 480
Db
421 wfrgttmyalyitvhgysflitflfgmvvvalvvmvkiftlrsatavkrgkrkvvltll 480
QY 481 GLSSLVGVTWGLAIFTPLGLSTVYIFALFNSLQG 514
Db
481 glsslvgvtwglaitfplglstvyifalfnsiqg 514

RESULT 3

AAE04449

ID AAE04449 standard; Protein; 508 AA.

XX AC AAE04449;

XX DT 04-SEP-2001 (first entry)

XX DE Human novel G-protein coupled receptor (NGPCR) protein #2.

XX KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
XX KW drug screening; transmembrane protein; mental disorder; gene therapy;
XX KW nootropic; pharmacogenomics; behavioural disorder.

XX OS Homo sapiens.

XX PN W0200142287-A2.

XX PD 14-JUN-2001.

XX PF 07-DEC-2000; 2000WO-US33241.

XX PR 07-DEC-1999; 99US-0169427.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;

XX PI Sands AT;

XX DR WPI; 2001-381634/40.

XX DR N-PSDB; AAD08759.

XX PT Novel polynucleotides encoding human G protein coupled receptors useful
XX PT for drug screening, diagnosis and in gene therapy of physiological or
XX PT behavioural disorders -

XX PS Claim 4; Page 62-63; 91pp; English.

XX CC The present sequence is human novel G protein coupled receptor
XX CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
XX CC cellular membrane and are involved in signal transduction after ligand

CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
XX related to NGPCR.

SQ Sequence 508 AA;

Query Match 88.5%; Score 2548; DB 22; Length 508;

Best Local Similarity 99.2%; Pred. No. 2.5e-255;

Matches 486; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 MATPRGLGALLLLPTSGOEKPTGPRNCLGNSNMNMFNLNDKALCFKCRQSGSD 60

Db 1 matprglgalllllptsggkptgprntclgnsnmnydfnlndkalcftkcrqsgsd 60

QY 61 SCNENLQRYWLYNEAHLMEGLTKQVNTPELKVQLVQLNSTNTAEDEFYSLEPSQVPRQV 120

Db 61 scnlenlqrywlyneahlmkegltqkntpfkalkvnlstaeafyfslepsqvprqv 120

QY 121 MKDEKPPDRVRLPKSLFRSLPGNRSVVRVLAVTILDIGPGTLFGKPRGLGDSGVLLNR 180

Db 121 mkdedkppdrvr.lpkslfrslpgnrsvvr.lavtildigpgtl.fkprglgdsgvllnr 180

QY 181 LVGLSVGQHVHTKLAEPLEIVFSHQRPNNMTLTCVFWDVTKGTTGDMSSGCSFEVRPE 240

Db 181 lvglsvghvhtklaepleivfshqrpnnmtlctcvfwdvktgtdmssescstevrpe 240

QY 241 GTVCCDHLTFFALLRPTLDQSTVHLTRISQAGCGVSMIFLAFTIILYAFRLSRERF 300

Db 241 gtvcccdhlthfllrptldgstvhltrisoagcgvsmlfraftiilyafrlsrerf 300

QY 301 KSEDAPKIHVALGGSFLNLNLAFLVNVGSGSGSDAACWARGAVHFYLLCAFTWMGLEA 360

Db 301 ksedapkihnvalggsflnlalnlaflvnnvsgsgsdacwargavhfylcaftwmglea 360

QY 361 FHLYLLAVRVNTYFGHYFLKLSLVGGLPALMVIGTGSANSYGLYTIIRDRENTSLELC 420

Db 361 fhlyllavrvntyfghyflklslvgglpalmvigtgsansyglytirdrenrtslelc 420

QY 421 WFRGTTMYALYITVHGYSFLITFLFGMVVVALVVMVKIFTLRSATAVKRGKRRKVVLTLL 480

Db 421 wfrgttmyalyitvhgysflitflfgmvvvalvvmvkiftlrsatavkrgkrkvvltll 480

QY 481 GLSSLVGVTW 490

Db 481 glsslv--asw 488

RESULT 4

AAE04464

ID AAE04464 standard; Protein; 485 AA.

XX AC AAE04464;

XX DT 04-SEP-2001 (first entry)

XX DE Human novel G-protein coupled receptor (NGPCR) protein #17.

Human: novel G-protein coupled receptor; NGPCR; signal transduction; drug screening; transmembrane protein; mental disorder; gene therapy; nootropic; pharmacogenomics; behavioural disorder.

Homo sapiens.

WO200142287-A2.

14-JUN-2001.

07-DEC-2000; 2000WO-US33241.

07-DEC-1999; 99US-0169427.

(LEXI-) LEXICON GENETICS INC.

Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B; Sands AT;

WPI: 2001-381634/40.

N-PSDB; AAD08774.

Novel polynucleotides encoding human G protein coupled receptors useful for drug screening, diagnosis and in gene therapy of physiological or behavioral disorders -

Claim 5; Page 79-80; 91pp; English.

The present sequence is human novel G protein coupled receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span the cellular membrane and are involved in signal transduction after ligand binding. NGPCR DNA sequences are useful as hybridisation probes for screening libraries and assessing gene expression patterns. Sequences derived from regions adjacent to the intron/exon boundaries of NGPCR gene can be used to design primers for use in amplification assays to detect mutations within the exons, splice sites, introns that can be used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene products are useful for identifying compounds that modulate gene expression or gene product activity. Such compounds are useful in the treatment of mental, physiological or behavioural disorders and diseases. NGPCR DNA sequences and antibodies are useful for diagnostic and prognostic evaluation of disorders related to NGPCR function and for the identification of subjects having a predisposition to such disorders. NGPCR DNA sequences are also useful for drug screening and in gene therapy for modulating NGPCR expression and to produce genetically engineered host cells to express NGPCR products in vivo. The encoded NGPCR proteins are useful for generating antibodies, as reagents in diagnostic assays and for identifying other cellular gene products related to NGPCR.

Sequence 485 AA;

Query Match
Best Local Similarity 86.7%; Score 2494; DB 22; Length 485;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATPRGIGALLLLLLLTSQGRTEGPRNTCLGSNNWYIFNNDKALCFTKCRSGSD 60
DB 1 matprgigalllllllptsggkptegprntclgsgnnmydifnldkalcftkcrsgsd 60

QY 61 SCNVENLQRYWLYEAHLMKEGLTKQVNTPTFLKALVQNLSTNAEDPFYSLEPSQVPRQV 120
DB 61 scnvenlqrywlyeahlmkegltkqvntptflkalvqnlstnaedpfyslepsqvprqv 120

QY 121 MKDEKPPDRVRLPKSLFRSLPGRNSRVRLAVTILDTGPGTFLKPRGLGDSGVYLNRR 180
DB 121 mkdekppdrvrpkpsfrslpgrnsrvrlavtildtgp tflkprglgdsqylnrr 180

QY 181 LVGLSVQCMHVTKLAEPLEIVFSHQRPNNMTITCVFMDWTKGTGDWSSEGSGTEVRPE 240
DB 181 lvglsvqcmhvtklaeplevfshqrpnnmtitcvfmdwtkgtgdwssegsgtevrpe 240

QY 241 GTVCCDHLTFPALLRPTLDQSTVHILTRISQAGCGVSMIFLAFTIILYAFRLSRERF 300
DB 241 gtvccdhltffalllrptldqstvhiltrisqagcgvsmiflaftiilyafrlsrerf 300

QY 301 KSEDAPKIHVALGGSFLFLNLAFVLNVVSGSGSDAACWARGAVPHYFLLCFTWMLGA 360
DB 301 ksedapkihvalggslfllnlafvlnvvgsgsgdaacwargavhyfllcftwmlga 360

QY 361 FHLYLLAVRVFNTYFGHYFLKLSLVGWLPAVMVITGTSANSYGLYTTIRDRNRTSLEIC 420
DB 361 fhlyllavrvfntyfghyflklsylvgwlpalvmvigtgsansyglytirdrenrtsleic 420

QY 421 WPREGTTMVALYTVHGYPFLITFLFGMVVVALVWVKIFTLSRATAVKERGNR 473
DB 421 wpregttmvalyitvhgyflitflfgmvvvalvwwkiftlsratavkergnr 473

RESULT 5
AAE04454
ID AAE04454 standard; Protein; 504 AA.
XX
AC AAE04454;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) protein #7.
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX
OS Homo sapiens.
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI: 2001-381634/40.
DR N-PSDB; AAD08764.
XX
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders -
XX
PS Disclosure; Page 68-69; 91pp; English.
XX
CC The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC products are useful for generating antibodies and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded

CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 504 AA;

Query Match 86.7%; Score 2494; DB 22; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-249;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATPRGLGALLLLPTSGQKPTGPRNTCLGNNMYDIFNLNDKALCFKCRQSGSD 60
DB 1 MATPRGLGALLLLPTSGQKPTGPRNTCLGNNMYDIFNLNDKALCFKCRQSGSD 60
QY 61 SCNVENLQRYWLNIEAHLMEKGLTQKNTPFLLKALVQNLSTNTAEDFFSLSPSQVPRQV 120
DB 61 SCNVENLQRYWLNIEAHLMEKGLTQKNTPFLLKALVQNLSTNTAEDFFSLSPSQVPRQV 120
QY 121 MKDEDKPPDRVRLPKSLFRSLPCNRSVVRKLVATIIDIGPTLFKGPRLGLDGGSGVLNRR 180
DB 121 MKDEDKPPDRVRLPKSLFRSLPCNRSVVRKLVATIIDIGPTLFKGPRLGLDGGSGVLNRR 180
QY 181 LVGLSVGOMHVTKLAEPLEIVESHORPPNMTLTCVFDVTKGTTGDSSECCSTEVRPE 240
DB 181 LVGLSVGOMHVTKLAEPLEIVESHORPPNMTLTCVFDVTKGTTGDSSECCSTEVRPE 240
QY 241 GTVCCDHLTFFALLRLPTDQSTVHILTRISQAGCGVSMIFLAFTIILYAFRLSRERF 300
DB 241 GTVCCDHLTFFALLRLPTDQSTVHILTRISQAGCGVSMIFLAFTIILYAFRLSRERF 300
QY 301 KSEDAPKIHVALGGSFLNLNLAFLVNVGSGKGSDAACWARGAVPHYLLCAFTWMGLEA 360
DB 301 KSEDAPKIHVALGGSFLNLNLAFLVNVGSGKGSDAACWARGAVPHYLLCAFTWMGLEA 360
QY 361 FHLYLAVRVNTYFGHYFLKLSLVGWGLPALMVIGTGSANSYGLYTRDRNRTSLBLC 420
DB 361 FHLYLAVRVNTYFGHYFLKLSLVGWGLPALMVIGTGSANSYGLYTRDRNRTSLBLC 420
QY 421 WREGTMYALYITVHGFIIFLFGMVVVALVVMKIFTLSRATAVKRGKNR 473
DB 421 WREGTMYALYITVHGFIIFLFGMVVVALVVMKIFTLSRATAVKRGKNR 473

RESULT 6
AAE04471
ID AAE04471 standard; Protein; 400 AA.
AC AAE04471;
XX
DT 04-SEP-2001 (first entry)
DE Human novel G-protein coupled receptor (NGPCR) protein #24.
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX
OS Homo sapiens.
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 95US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Doncho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX

DR WPI; 2001-381634/40.
XX N-PSDB; AAD08781.
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders -
XX
XX Disclosure; Page 87-88; 91pp; English.
XX
CC The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 400 AA;

Query Match 70.3%; Score 2022; DB 22; Length 400;
Best Local Similarity 99.5%; Pred. No. 7.3e-201;
Matches 386; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 162 LFKGPRGLGDSGVNLRNLVGLSVGOMHVTKLAEPLEIVESHORPPNMTLTCVFDVDT 221
DB 13 LSGQPRGLGDSGVNLRNLVGLSVGOMHVTKLAEPLEIVESHORPPNMTLTCVFDVDT 72
QY 222 KGTGDSSECCSTEVRPEGTVCDDHLTFFALLRLPTDQSTVHILTRISQAGCGVSMI 281
DB 73 KGTGDSSECCSTEVRPEGTVCDDHLTFFALLRLPTDQSTVHILTRISQAGCGVSMI 132
QY 282 FLAFTIILYAFRLSRERFSEKEDAPKIHVALGGSFLNLNLAFLVNVGSGKGSDAACWAR 341
DB 133 FLAFTIILYAFRLSRERFSEKEDAPKIHVALGGSFLNLNLAFLVNVGSGKGSDAACWAR 192
QY 342 GAVPHYLLCAFTWMGLEAFHLYLLAVRVNTYFGHYFLKLSLVGWGLPALMVIGTGSAN 401
DB 193 GAVPHYLLCAFTWMGLEAFHLYLLAVRVNTYFGHYFLKLSLVGWGLPALMVIGTGSAN 252
QY 402 SYGLYTRDRNRTSLBLCWFRGTTMYALYITVHGFIIFLFGMVVVALVVMKIFTL 461
DB 253 SYGLYTRDRNRTSLBLCWFRGTTMYALYITVHGFIIFLFGMVVVALVVMKIFTL 312
QY 462 RATAVKRGKNRKVLTLGLSSILVGVTVWGLAIFTPLGLSTVYIFALFNSLQGVFCCWF 521
DB 313 RATAVKRGKNRKVLTLGLSSILVGVTVWGLAIFTPLGLSTVYIFALFNSLQGVFCCWF 372
QY 522 TILYLPQSQTTSVSSSTARLDQAHSASOE 549
DB 373 TILYLPQSQTTSVSSSTARLDQAHSASOE 400

RESULT 7
AAE04461
ID AAE04461 standard; Protein; 372 AA.
XX
AC AAE04461;
XX

DT 04-SEP-2001 (first entry)
 XX Human novel G-protein coupled receptor (NGPCR) protein #14.
 DE
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 XX drug screening; transmembrane protein; mental disorder; gene therapy;
 KW neotropic; pharmacogenomics; behavioural disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200142287-A2.
 PN
 XX 14-JUN-2001.
 PD
 XX
 PF 07-DEC-2000; 2000WO-US33241.
 XX
 XX 07-DEC-1999; 99US-0169427.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX
 DR WPI; 2001-381634/40.
 DR N-PSDB; AAD08771.
 XX
 XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioural disorders -
 XX
 PS Disclosure; Page 76-77; 91pp; English.
 XX
 CC The present sequence is human novel G protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.
 XX
 XX Sequence 372 AA;
 SQ

Query Match 63.9%; Score 1840; DB 22; Length 372;
 Best Local Similarity 99.4%; Pred. No. 5.1e-182;
 Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 162 LFKGPRGLGDSGLNRLVGLSVGMVHTKLAEPLEIVFSHQRPPTNLTCVFDVDT 221
 | :|||||
 Db 13 lsgqprlgdsgvlnrlvglsvgmhvklaepieivfshqrpptmtlcvfdvdt 72
 | :|||||
 Qy 222 KGTGDSSECSCEVRPEGVCCDHLTFEALLRPTLDQSTVHILTRISQAGCGVSMI 281
 | :|||||
 Db 73 kgtgdswseccstevrpegvtcccdhltfllrptldqstvhiltrisqgcvsmi 132
 | :|||||
 Qy 282 FLATILYAFRLRSERFKSEDPKTHVALGSLFLNLAFILNVNKGSGKSDAACWAR 341
 | :|||||
 Db 133 flatililyafirlsrerfkseapkhvalgslflnlafilnvngsgsgksdaacwar 192
 | :|||||

QY 342 GAVFHYELLCAFTWGMGLEAFHLYLLAVRVNTVGHVFLKLSLVGWLPAVMVIGTGSAN 401
 | :|||||
 Db 193 gavfhyellcaftwgmgleafhlyllavrvntvghvflklslvgwlpalvmvigtgsan 252
 | :|||||
 QY 402 SYGLYIRDRNRTSLELCWFREGTMYALYITVHGYFLITFLFGMVVIALVVKIFTLS 461
 | :|||||
 Db 253 syglyirdrentslelcwfregrtmyalyitvhgyflitflfgmvvialvvkiftls 312
 | :|||||
 QY 462 RATAVKERGKNRKKVLTLLGLSLVGVWTWGLAIFTPLGLSTVYIFALFNSLOG 514
 | :|||||
 Db 313 ratavkergknrkkvltllglsslvgvtwglaitfplgltstvyifalfnslsg 365
 | :|||||

RESULT 8
 AAE04470
 ID AAE04470 standard; Protein; 339 AA.
 XX
 AC AAE04470;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) protein #23.
 XX
 KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 KW drug screening; transmembrane protein; mental disorder; gene therapy;
 KW neotropic; pharmacogenomics; behavioural disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200142287-A2.
 PN
 XX 14-JUN-2001.
 PD
 XX
 PF 07-DEC-2000; 2000WO-US33241.
 XX
 XX 07-DEC-1999; 99US-0169427.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AT;
 XX
 DR WPI; 2001-381634/40.
 DR N-PSDB; AAD08780.
 XX
 XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioural disorders -
 XX
 PS Disclosure; Page 86-87; 91pp; English.
 XX
 CC The present sequence is human novel G protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.
 XX


```
SQ Sequence 339 AA;
Query Match 61.8%; Score 1778; DB 22; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.2e-175;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 MTLTCVFDVTKGTTGDWSEGCSTEVPRPEGVCCDHLTPFALLRPTLDQSTVHILTR 270
DQ 1 mlttcvfdvtkgtgdwssegcstevprpegtvcccdhltpfallrptldgstvhilr 60
QY 271 ISGAGCGVMIFLAFTIILYAFILRSRERFKSEDPKIHVALGGSFLNLALFLNVGSG 330
DQ 61 isagcgvmiflaftiilyafilrsrerfksepkihvalggsflnlalflnvsgs 120
QY 331 SKGSDAACWARGAVFHYLLCAFTWMLGFAFHLYLAVRVNTYFGHYFLKLSLVGWL 390
DQ 121 skgsdaacwargavfhyllcaftwmlgfaafhllylavrvntfyghyflklslvgwlp 180
QY 391 ALMVIGTGSANSGLYTIIRDRENTSLELCWFREGTMYALYTVHGYFLITFLGMYVL 450
DQ 181 almvigtgsansglytiirdrenrtslelcwfiregtmyalytvhgyflitflgmvl 240
QY 451 ALVVMKIFTLRSRATVAKRGKNNKVVTLGLSSLVGVTWGLAIFTPGLSTVYIFALFN 510
DQ 241 alvwmkiftlrsratavakrgknkvvltlglsslvgvwtwglaitfpglstvyifaln 300
QY 511 SLOGVFICWFTILYIPSOSTTVSSSTARLDQAHSASQE 549
DQ 301 slggvficwftilyipsgsttvssstarldqahsasqe 339

RESULT 9
AAE04451
ID AAE04451 standard; Protein; 359 AA.
XX
AC AAE04451;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) protein #4.
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX
OS Homo sapiens.
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
PI WPI: 2001-381634/40.
XX
DR N-PSDB; AAD08761.
XX
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioural disorders -
XX
PS Disclosure; Page 65-66; 91pp; English.
XX
CC The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
```

```
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products.
XX
SQ Sequence 359 AA;
Query Match 58.8%; Score 1692; DB 22; Length 359;
Best Local Similarity 98.2%; Pred. No. 1.1e-166;
Matches 323; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 162 LFKGPRGLGDSGVNRLVGLSVGQMHVTKLAEPLEIVFESHQRPPTTLTCVFDVDT 221
DQ 13 lskgprglgdsgvlnrlvglsvgmhvtklaepileivfeshqrppttltcvfdvdt.72
QY 222 KGTGDSSEGCSTEVPRPEGVCCDHLTPFALLRPTLDQSTVHILTRISQAGCGVSMI 281
DQ 73 kgtgdwssegcstevprpegtvcccdhltpfallrptldgstvhiltrisagcgvsmi 132
QY 282 FLAFTIILYAFILRSRERFKSEDPKIHVALGGSFLNLALFLNVGSGSGSDAACWARG 341
DQ 133 flaftiilyafilrsrerfksepkihvalggsflnlalflnvsgsgsdacwarg 192
QY 342 GAVHYFLLCFAFTWMLGFAFHLYLAVRVNTYFGHYFLKLSLVGWLPAIMVIGTGSAN 401
DQ 193 gavhyflclcaftwmlgfaafhllylavrvntfyghyflklslvgwlpaimvigtgsan 252
QY 402 SYGLYTIIRDRENTSLELCWFREGTMYALYTVHGYFLITFLGMYVLALVVMKIFTL 461
DQ 253 syglytiirdrenrtslelcwfiregtmyalytvhgyflitflgmvlalvwmkiftls 312
QY 462 RATAVKEGKNNKVVTLGLSSLVGVTW 490
DQ 313 ratavkegknrkvvltlglsslvgvtw --asw 339

RESULT 10
AAE04466
ID AAE04466 standard; Protein; 336 AA.
XX
AC AAE04466;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) protein #19.
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX
OS Homo sapiens.
XX
PN WO200142287-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
```

PR 07-DEC-1999; 99US-0169427.
XX (LEXI-) LEXICON GENETICS INC.
PA Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-381634/40.
DR N-PSDB; AAD08776.
XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders
XX Disclosure: Page 82; 91pp; English.
PS
XX The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 336 AA;

Query Match 56.9%; Score 1638; DB 22; Length 336;
Best Local Similarity 99.4%; Pred. No. 4.1e-161;
Matches 310; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 162 LFKGPRGLGDSGVNLRNLVGLSVGQMHVTKLAEPLEIVFSSHQRPNNMTLTCVFDVDT 221
Db | : |||||
Qy 13 lsqgprlgldgsgvlnrlvlgsvgmhvtklaepleivfshqrpnnmtltcvfdvdt 72
222 KGTGDMSSGCGSTEVRPEGTVCCEHLLTFALLRPTLDQSTVHILTRISQAGCGVSMI 281
Db |||||
Qy 73 kgtgdwssegcstevrpegtvcccdhlftfallrptldgstvhiltrisqagcgvsml 132
282 FLAFTIILYAFRLSRERFKSEDAPIKHVALGGSFLNLAFNLVNVGSGSGSDAACWAR 341
Db |||||
Qy 133 flaftiilayafirlsrerfkseapdkhvalggsflnlafnlvngsgsgsdacwar 192
342 GAVFHYELLCRAFTWGLEAHPHLYLLAVRVNTFYGHYFLKLSLVGGLPALMVIPTGSAN 401
Db |||||
Qy 193 gavfhyflcaftwmgleahphlyllavrvntfyghyflklsilvgglpalmvigtgsan 252
402 SYGLTYTIRDRENTSLELCWFREGTQWYALYITVGHYFLTELFGMVVLAIVVVKFTLS 461
Db |||||
Qy 253 syglytirdrenrtslslcwfregtqwmalyitvghyfltelifgmvvlaivvkvkftls 312
462 RATAVKERGNR 473
Db |||||
313 ratavkergnr 324

RESULT 11
AAE04456
ID AAE04456 standard; Protein; 355 AA.

XX AAE04456;
AC 04-SEP-2001 (first entry)
XX Human novel G-protein coupled receptor (NGPCR) protein #9.
DT
XX
DE
XX
KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX
OS Homo sapiens.
XX WO200142287-A2.
PN
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-US33241.
XX
XX 07-DEC-1999; 99US-0169427.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
XX WPI: 2001-381634/40.
DR N-PSDB; AAD08766.
XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders -
XX
XX Disclosure: Page 70; 91pp; English.
XX
XX The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 355 AA;

Query Match 56.9%; Score 1638; DB 22; Length 355;
Best Local Similarity 99.4%; Pred. No. 4.5e-161;
Matches 310; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 162 LFKGPRGLGDSGVNLRNLVGLSVGQMHVTKLAEPLEIVFSSHQRPNNMTLTCVFDVDT 221
Db | : |||||
Qy 13 lsqgprlgldgsgvlnrlvlgsvgmhvtklaepleivfshqrpnnmtltcvfdvdt 72
222 KGTGDMSSGCGSTEVRPEGTVCCEHLLTFALLRPTLDQSTVHILTRISQAGCGVSMI 281
Db |||||
Qy 73 kgtgdwssegcstevrpegtvcccdhlftfallrptldgstvhiltrisqagcgvsml 132
282 FLAFTIILYAFRLSRERFKSEDAPIKHVALGGSFLNLAFNLVNVGSGSGSDAACWAR 341

|||||
Db 133 flattilyafllrsrefksedapkihaigslflnlnlaflnvvgsgskgsdaacwar 192
QY 342 GAVFHYELLCFAFTWMLGAEAFHLYLLAVRVNTYFGHYFLKLSLVGWLGPALMVGIGTSAN 401
Db 193 gavfhyflcaftwmgleahlyllavrvntfyghyflklslvgwglpalmvigtgsan 252
QY 402 SYGLYTTIRDRENTSLELCWREGTMYALXIIVHGYFLITFLFGWVVALVWKIFTLIS 461
Db 253 sygliyrtirdrenrtslelcwregtmyalyitvhgyflitflfgmvvialvwwkiftis 312
QY 462 RATAVKRGKNR 473
Db 313 ratavkergknr 324

RESULT 12
AAE04460
ID AAE04460 standard; Protein; 311 AA.
XX AC AAE04460;
XX DT 04-SEP-2001 (first entry)
XX DE Human novel G-protein coupled receptor (NGPCR) protein #13.
XX KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX OS Homo sapiens.
XX PN WO200142287-A2.
XX PD 14-JUN-2001.
XX PF 07-DEC-2000; 2000WO-US33241.
XX PR 07-DEC-1999; 99US-0169427.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
XX Sands AT;
XX WPI: 2001-381634/40.
XX DR N-PSDB; AAD08770.
XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders
XX Disclosure; Page 75-76; 91pp; English.

The present sequence is human novel G protein coupled receptor
(NGPCR) protein. NGPCRs are transmembrane proteins that span the
cellular membrane and are involved in signal transduction after ligand
binding. NGPCR DNA sequences are useful as hybridisation probes for
screening libraries and assessing gene expression patterns. Sequences
derived from regions adjacent to the intron/exon boundaries of NGPCR
gene can be used to design primers for use in amplification assays to
detect mutations within the exons, splice sites, introns that can be
used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
products are useful for identifying compounds that modulate gene
expression or gene product activity. Such compounds are useful in the
treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in

CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX SQ Sequence 311 AA;
Query Match 55.5%; Score 1596; DB 22; Length 311;
Best Local Similarity 100.0%; Pred. No. 8.4e-157;
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 MLTCTCFVDTVKTGCDNSSECSSTEVREPEGVCCDHLTFPALLRLPTLDQSTVHILTR 270
Db 1 mltctcfvdtvktgtdwssegcstevrpegtvcccdhltfallrlpctldqstvhiltr 60
QY 271 ISOAGCGVSMIFLAFTIILYAFRLSRERFKSEDAPIKHVALGGSLFLINLAFLVNVS 330
Db 61 isqagcgvsmiflaftiilayafllrsrefksedapkihaigslflnlnlaflnvvgsg 120
QY 331 SKGSDAACWARGAVFHYFLLCAPTWMGLEAFHLYLLAVRVNTYFGHYFLKLSLVGWL 390
Db 121 skgsdaacwargavfhyflcaftwmgleahlyllavrvntfyghyflklslvgwglp 180
QY 391 ALMVIGTGSANSGLYTIIRDRENTSLELCWREGTMYALXIIVHGYFLITFLFGMVVL 450
Db 181 almvigtgsansglytiirdrenrtslelcwregtmyalyitvhgyflitflfgmvvl 240
QY 451 ALVWVKIFTLRSATAVKRGKNRKVVLTILGLSSLVGVTWGLAIFTPLGLSTVIYIFALFN 510
Db 241 alvwvkiftlrsatavkergknrkvvltlilgsslvgtwglaitftplglstviyifalfn 300
QY 511 SLOG 514
Db 301 slgg 304

RESULT 13
AAE04450
ID AAE04450 standard; Protein; 298 AA.
XX AC AAE04450;
XX DT 04-SEP-2001 (first entry)
XX DE Human novel G-protein coupled receptor (NGPCR) protein #3.
XX KW Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX OS Homo sapiens.
XX PN WO200142287-A2.
XX PD 14-JUN-2001.
XX PF 07-DEC-2000; 2000WO-US33241.
XX PR 07-DEC-1999; 99US-0169427.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
XX Sands AT;
XX WPI: 2001-381634/40.
XX DR N-PSDB; AAD08760.
XX Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders
XX Disclosure; Page 64; 91pp; English.

CC The present sequence is human novel G protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.

XX Sequence 298 AA;

Query Match 50.3%; Score 1448; DB 22; Length 298;
 Best Local Similarity 98.6%; Pred. No. 1.8e-141;
 Matches 276; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy 211 MTLTCVFDVTKGTGGWSEGCSTVRPEGVCCDHLTFALLRPTLDQSTVHLTR 270
 Db 1 mtltcvfdvktgtdwsegcstevrpegtvcccdhltfallrptldgstvhltr 60
 Qy 271 ISOAGCGVSMIFLAFTIILYAFRLSRERFKSEDPKIHVALGGSFLNLNLAFLVNVGSG 330
 Db 61 isqagcgvmiflaftiilyafirlsrerfkseadapkihvalggsflnlaflnvngsg 120
 Qy 331 SKGSDAACHWARGAVHYFLLCFTWGLFAFHLVLLAVRVNTYFGHYFLKLSLVGWLGP 390
 Db 121 skgsdaacwargavhyflcftwmgleaflhlyllavrvntfyghyflklslvgwglp 180
 Qy 391 ALMVIGTGSANSYGLYTIIRDRENTSLELCWFREGTMYALYITVHGYSFLITFLFGMVVL 450
 Db 181 almvigtgsansyglytirdrenrtselelcwfregrtmyalyitvhgyflitflfgmvvl 240
 Qy 451 ALVWKIFTLRSATAVKRGKRVLTLLGLSSLGVGVTVW 490
 Db 241 alvwwkiftlsratavkrgknrvltllgssl--asw 278

RESULT 14

AAE04465
 ID AAE04465 standard; Protein; 275 AA.

XX AAE04465;

XX 04-SEP-2001 (first entry)

XX Human novel G-protein coupled receptor (NGPCR) protein #18.

XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
 XX drug screening; transmembrane protein; mental disorder; gene therapy;
 XX nootropic; pharmacogenomics; behavioural disorder.

XX Homo sapiens.

XX WO200142287-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-US33241.

XX 07-DEC-1999; 99US-0169427.

XX

PA (LEXI-) LEXICON GENETICS INC.

PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Sands AN;

XX WPI; 2001-381634/40.

DR N-PSDB; AAD08775.

XX Novel polynucleotides encoding human G protein coupled receptors useful
 PT for drug screening, diagnosis and in gene therapy of physiological or
 PT behavioural disorders

XX Disclosure; Page 81; 91pp; English.

XX The present sequence is human novel G protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
 CC cellular membrane and are involved in signal transduction after ligand
 CC binding. NGPCR DNA sequences are useful as hybridisation probes for
 CC screening libraries and assessing gene expression patterns. Sequences
 CC derived from regions adjacent to the intron/exon boundaries of NGPCR
 CC gene can be used to design primers for use in amplification assays to
 CC detect mutations within the exons, splice sites, introns that can be
 CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
 CC products are useful for identifying compounds that modulate gene
 CC expression or gene product activity. Such compounds are useful in the
 CC treatment of mental, physiological or behavioural disorders and diseases.
 CC NGPCR DNA sequences and antibodies are useful for diagnostic and
 CC prognostic evaluation of disorders related to NGPCR function and for the
 CC identification of subjects having a predisposition to such disorders.
 CC NGPCR DNA sequences are also useful for drug screening and in gene
 CC therapy for modulating NGPCR expression and to produce genetically
 CC engineered host cells to express NGPCR products in vivo. The encoded
 CC NGPCR proteins are useful for generating antibodies, as reagents in
 CC diagnostic assays and for identifying other cellular gene products
 CC related to NGPCR.

XX Sequence 275 AA;

Query Match 48.4%; Score 1394; DB 22; Length 275;
 Best Local Similarity 100.0%; Pred. No. 6.6e-136;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 MTLTCVFDVTKGTGGWSEGCSTVRPEGVCCDHLTFALLRPTLDQSTVHLTR 270

Db 1 mtltcvfdvktgtdwsegcstevrpegtvcccdhltfallrptldgstvhltr 60

Qy 271 ISOAGCGVSMIFLAFTIILYAFRLSRERFKSEDPKIHVALGGSFLNLNLAFLVNVGSG 330

Db 61 isqagcgvmiflaftiilyafirlsrerfkseadapkihvalggsflnlaflnvngsg 120

Qy 331 SKGSDAACHWARGAVHYFLLCFTWGLFAFHLVLLAVRVNTYFGHYFLKLSLVGWLGP 390

Db 121 skgsdaacwargavhyflcftwmgleaflhlyllavrvntfyghyflklslvgwglp 180

Qy 391 ALMVIGTGSANSYGLYTIIRDRENTSLELCWFREGTMYALYITVHGYSFLITFLFGMVVL 450

Db 181 almvigtgsansyglytirdrenrtselelcwfregrtmyalyitvhgyflitflfgmvvl 240

Qy 451 ALVWKIFTLRSATAVKRGKRVLTLLGLSSLGVGVTVW 473

Db 241 alvwwkiftlsratavkrgknrvltllgssl--asw 263

RESULT 15

AAE04455
 ID AAE04455 standard; Protein; 294 AA.

XX AAE04455;

XX 04-SEP-2001 (first entry)

XX

DE Human novel G-protein coupled receptor (NGPCR) protein #8.
XX Human; novel G-protein coupled receptor; NGPCR; signal transduction;
KW drug screening; transmembrane protein; mental disorder; gene therapy;
KW nootropic; pharmacogenomics; behavioural disorder.
XX Homo sapiens.
OS
XX WO200142287-A2.
XX 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33241.
XX
PR 07-DEC-1999; 99US-0169427.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI: 2001-381634/40.
DR N-PSDB: AAD08765.
XX
PT Novel polynucleotides encoding human G protein coupled receptors useful
PT for drug screening, diagnosis and in gene therapy of physiological or
PT behavioral disorders
XX
PS Disclosure; Page 69-70; 9lpp; English.
XX

CC The present sequence is human novel G protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span the
CC cellular membrane and are involved in signal transduction after ligand
CC binding. NGPCR DNA sequences are useful as hybridisation probes for
CC screening libraries and assessing gene expression patterns. Sequences
CC derived from regions adjacent to the intron/exon boundaries of NGPCR
CC gene can be used to design primers for use in amplification assays to
CC detect mutations within the exons, splice sites, introns that can be
CC used in diagnostics and pharmacogenomics. NGPCR sequences and/or gene
CC products are useful for identifying compounds that modulate gene
CC expression or gene product activity. Such compounds are useful in the
CC treatment of mental, physiological or behavioural disorders and diseases.
CC NGPCR DNA sequences and antibodies are useful for diagnostic and
CC prognostic evaluation of disorders related to NGPCR function and for the
CC identification of subjects having a predisposition to such disorders.
CC NGPCR DNA sequences are also useful for drug screening and in gene
CC therapy for modulating NGPCR expression and to produce genetically
CC engineered host cells to express NGPCR products in vivo. The encoded
CC NGPCR proteins are useful for generating antibodies, as reagents in
CC diagnostic assays and for identifying other cellular gene products
CC related to NGPCR.
XX
SQ Sequence 294 AA;

Query Match 48.4%; Score 1394; DB 22; Length 294;
Best Local Similarity 100.0%; Pred. No. 7.3e-136;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 MTLTCVFDVTKGTGWSSEGGSTEVPRPGTVCDDHLTFALLRPTLDQSTVHLTR 270
DB 1 mtlctvfdvdkgtgdwssegcstevprpgtvcddhlftfallrptldqstvhiltr 60
QY 271 ISOAGCGVSMTFLAFTIILYAFILSLRERFKSEDAKIHVALGSLFLNLAFLVNMGSC 330
DB 61 isqagcgvsmflaftlilyafilrsrfrkdsedapkihvalgsgsflnlaflvnvgsg 120
QY 331 SKGSDAACWARGAVFHYFLLCFTWGLFAFLYLLAVRVNTYFGHYFLKLSLVGWLGP 390
DB 121 skgsdaacwargavfhyfllicftwmgleafhlyllavrvncyfghyflksilvgwglp 180
QY 391 ALMWIGTGSANSGLYITRDRENTSLELCWFREGTMTYALYITVHGFLITFLFGMVYL 450

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 21:31:23 ; Search time 24.88 seconds
(without alignments)
538.973 Million cell updates/sec

Title: US-09-733-387-44
Perfect score: 2878
Sequence: 1 MATPRGLGALLLLLLLPTSG.....STTVSSSTARLDQASASQE 549

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTIUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	316	11.0	521	US-08-956-322-4	Sequence 4, Appli
2	316	11.0	652	US-09-110-116-1	Sequence 1, Appli
3	316	11.0	652	US-08-956-322-2	Sequence 2, Appli
4	300	10.4	344	US-09-110-116-4	Sequence 4, Appli
5	292.5	10.2	1052	US-08-852-806-2	Sequence 2, Appli
6	292.5	10.2	1052	US-09-163-669-2	Sequence 2, Appli
7	264	9.2	240	US-09-370-098-5	Sequence 5, Appli
8	254.5	8.8	886	US-09-110-116-3	Sequence 3, Appli
9	213	7.4	884	US-08-465-976A-2	Sequence 2, Appli
10	213	7.4	884	US-08-982-412-2	Sequence 2, Appli
11	199	6.9	235	US-09-370-098-6	Sequence 6, Appli
12	198.5	6.9	231	US-09-370-098-4	Sequence 4, Appli
13	177.5	6.2	415	US-08-110-286A-2	Sequence 2, Appli
14	177.5	6.2	431	US-08-981-189B-11	Sequence 11, Appli
15	174.5	6.1	431	US-08-381-433A-2	Sequence 2, Appli
16	173.5	6.0	411	US-08-381-433A-4	Sequence 4, Appli
17	173.5	6.0	411	US-08-981-189B-12	Sequence 12, Appli
18	173.5	6.0	431	US-08-981-189B-13	Sequence 13, Appli
19	172.5	6.0	415	US-08-110-286A-6	Sequence 6, Appli
20	172.5	6.0	415	US-08-981-189B-10	Sequence 10, Appli
21	158	5.5	509	US-08-845-546-2	Sequence 2, Appli
22	154.5	5.4	411	US-08-381-433A-8	Sequence 8, Appli
23	151	5.2	514	US-09-370-098-2	Sequence 2, Appli
24	149.5	5.2	553	US-08-845-546-12	Sequence 12, Appli
25	148.5	5.2	515	US-08-468-249A-18	Sequence 18, Appli
26	148.5	5.2	585	US-08-142-439A-6	Sequence 6, Appli
27	148.5	5.2	585	US-08-142-551B-125	Sequence 125, App

28	148.5	5.2	585	2	US-08-869-477-6	Sequence 6, Appli
29	148.5	5.2	585	2	US-08-468-249A-19	Sequence 19, Appli
30	143	5.0	482	1	US-08-142-439A-7	Sequence 7, Appli
31	143	5.0	482	2	US-08-869-477-7	Sequence 7, Appli
32	143	5.0	541	3	US-08-468-011A-2	Sequence 2, Appli
33	143	5.0	541	4	US-09-236-468A-2	Sequence 2, Appli
34	143	5.0	541	5	PCT-US95-07085-2	Sequence 2, Appli
35	140	4.9	482	1	US-07-792-885A-1	Sequence 1, Appli
36	140	4.9	593	2	US-08-468-249A-21	Sequence 21, Appli
37	138	4.8	269	4	US-09-370-098-3	Sequence 3, Appli
38	134	4.7	437	2	US-08-538-816A-2	Sequence 2, Appli
39	134	4.7	437	2	US-09-076-651-2	Sequence 2, Appli
40	134	4.7	437	4	US-09-208-394-2	Sequence 2, Appli
41	133.5	4.6	472	5	PCT-US94-09235-2	Sequence 2, Appli
42	133.5	4.6	474	1	US-08-453-742-2	Sequence 2, Appli
43	133.5	4.6	474	1	US-08-454-464-2	Sequence 2, Appli
44	133.5	4.6	474	1	US-08-453-222-2	Sequence 2, Appli
45	133.5	4.6	474	1	US-08-452-802-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-956-322-4
; Sequence 4, Application US/08956322
; Patent No. 6277977
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: MAO, JOYCE
; TITLE OF INVENTION: CDNA CLONE HAPO167 THAT ENCODES
; TITLE OF INVENTION: A HUMAN 7-TRANSMEMBRANE RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,322
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,329
; FILING DATE: 11-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70075
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-956-322-4

Query Match 11.0%; Score 316; DB 4; Length 521;
Best Local Similarity 28.8%; Pred. No. 1.4e-23;
Matches 105; Conservative 55; Mismatches 165; Indels 42; Gaps 16;

QY	177	LNRLVGLSVGQMHVTKLAEPLEIVFESH--QRPPNNMTLTCVFDWVTGKTGDWSSBGC--S34
Db	134	LNQSVSAIGKPRNNVLSKSVTLTFQHKMTPTSTKFKVFCYWKST--GQGSQWSRDCFL192
QY	235	TEVRPCTGCCDHLTFEALLRPTLDOSTVHLLTRISQAGCGVSMIFLAFIILYAFUR294
Db	193	IHWKSHMTMCNCHLSFAVLMALT--SQBEDPVLTVITYVGLSVSLCLLALLFTFLCK255
QY	295	LSRERPKSEDAPKIHVALGSGFLNLAFLVNNGSGSKGSDAACWARGAVFHYFLICAPT354
Db	252	AIQNTSTS-----LHLQLSICLELAHLFL--VGIDRTEPKVLCSTIAGALHYLYLAFT304
QY	355	WMGLEAFHLYLLA--VRVNTVTFGHVFLKLSL--VGWGLPALMVGITGSGANSYGLYITRD410
Db	305	WMLEGVHLLFTARNLTVTNNYSSINRLMKWIMFPFVPGYGVPA--VTVAISAASPHLYGTAD363
QY	411	RENRTSLELCWFREGTMTYALXITVHYGFILTFLCGMVVALVWV---KITLSRATV466
Db	364	R-----CWLHDGGMWSFL---GPVCAIFSANLVILFVWLKRLKLSLNSVET412
QY	467	KERGNRKKVTLTLGLSSL--VGVTW--GLAIFTPLGLSTVYIFALFNLSIQGVFCICWFT522
Db	413	IQ----NTRMLAFKATAQLFILCTWCLGLQLGVPAAQVMAYLFTIINSIQGFET---FL465
QY	523	ILYLPQS529
Db	466	VYCLLS472

```

RESULT      2
US-09-110-116-1
; Sequence 1, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong
; APPLICANT: Cohan, Victoria L.
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: HUMAN EMRI-LIKE G PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: PF-0550 US
; CURRENT APPLICATION NUMBER: US/09/110,116
; CURRENT FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 652
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 429905, EOSINOT03
US-09-110-116-1

```

Query Match	11.0%	Score 316;	DB 3;	Length 652;
Best Local Similarity	28.6%	Pred. No. 2e-23;		
Matches 105:	Conservative	55;	Mismatches 165;	Indels 42; Gaps 16;

[illegible]

Qy	466	RENTSLELCWCFREGTMMYALYTVVIGYFLLITFLGMYVLAVVW---- <td>466</td>	466
Db	411	RENTSLELCWCFREGTMMYALYTVVIGYFLLITFLGMYVLAVVW---- <td>411</td>	411
Db	495	R-----CWLHLDQGFMSFL---GPVCAIFSANLVILFVFWILKRLSLNSEVST	543
Qy	467	KERGKRNKKVLTLLGLSSL--VGVTW--GLAIFTPLGLSTVYIFALFNSLQGVFICWFT	532
Db	544	IQ-----NTRMLAFKATAQLFGLCTWCLGLLQVGPAQVMAYLFTIINSLQGFFI---FL	596
Qy	523	ILVLPQS 529	
Db	597	VYCLLSO 603	

RESULT 3
 US-08-956-322-2
 ; Sequence 2, Application US/08956322
 ; Patent No. 627797
 ; GENERAL INFORMATION:
 ; APPLICANT: SATHE, GANESH
 ; APPLICANT: MAO, JOYCE
 ; TITLE OF INVENTION: CDNA CLONE HAPO167 THAT ENCODES
 ; TITLE OF INVENTION: A HUMAN 7-TRANSMEMBRANE RECEPTOR
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RATNER & PRESTIA
 ; STREET: P.O. BOX 980
 ; CITY: VALLEY FORGE
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19482
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: Fastseq for windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/956,322
 ; FILING DATE: 23-OCT-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/049,329
 ; FILING DATE: 11-JUN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PRESTIA, PAUL F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: GH-70075
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEFAX: 610-407-0701
 ; TELEX: 846169
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 652 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PS-08-956-322-2

Query Match 11.0%; Score 316; DB 4; Length 652;
Best Local Similarity 28.6%; Pred. No. 2e-23;
Matches 105; Conservative 55; Mismatches 165; Indels 42; Gaps 16;

QY	177	LNNRLVGLSVGQMHVTKLAEPLEIVESH-QRRPPNMTLTCVFDVDTGKTGGTGDWSSBGC-S	234
Db	265	LNQVSGAAITGPRNNVSLKSVLTTFQHVAKMTSTKKVFCVYKMS-GQGSQMSRDGCEL	323
QY	235	TEVREPTGVCDDHLTFPALLRPTLDQSTVHLTRISQAGCVSMIFLAFTIILVAFER	294
Db	324	IYVKNSTHMCNCHLSFAYLMALT-SOEDPVLVTYVGLSVSLCLLALAAITLFLCK	382

QY 295 LSRERKSDAPKIHVALGSLFLLNLAFVLNVGSGKSDAACWARGAVHYFLICAPT 354
Db 383 AIONTS-----LHLQSLCLFLAHLFL--VGIDRTEPKVLCIIAGALHYLYLAAPT 435
QY 355 WMGLEAFHYLYLA--VRVNTYFGHYFLKSL--VWGMLPALMVIGTGSANSYGLYTRD 410
Db 436 WMLLEGVHLFLTARNLTVVNYSSINRLMKWIMFPVGVGVA--VTVAISAASWPHLYGTAD 494
QY 411 RENRTSLELCWREGTMYALYTVHGYFLITFLGMVVALVWV----KFTLSRAVAV 466
Db 495 R-----CWLHLOGFMWSEL---GPVCAIFSANLVFLFVILFKKLSLSEYST 543
QY 467 KERKNNKVKVLTILGLSSL--VGVTW--GLAIFTPLGLSTVYIFALENSLOGVFCCWFT 522
Db 544 IQ-----NTRMLAFKATAQLFILGCTWCLGLLQVGAQVMAVLFVILNLSQGGFI--FL 596
QY 523 ILYLPQS 529
Db 597 VYCLLSQ 603
RESULT 4
US-09-110-116-4
; Sequence 4, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong
; APPLICANT: Cohan, Victoria L.
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: HUMAN EMRL-LIKE G PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: PF-0550 US
; CURRENT APPLICATION NUMBER: US/09/110,116
; CURRENT FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 344
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 2935597, GenBank
US-09-110-116-4

Query Match 10.4%; Score 300; DB 3; Length 344;
Best Local Similarity 29.4%; Pred. No. 3.2e-22;
Matches 95; Conservative 50; Mismatches 148; Indels 30; Gaps 12;

QY 215 CVFWDVTKGTGDWSEGCST-EVRPEGTVCCHLTFALLRPTLDQSTVHILTRISQ 273
Db 4 CVFWEHGQCGCHWATTCSTIGTRDTSTICRCTLSSFAVLM-AHYDVQEDPVLTVITY 62
QY 274 ACGGSMIFLAFTIILYAFRLSRERKSDAPKIHVALGSLFLLNLAFVLNVGSGSKG 333
Db 63 MGLSVSLCLLLAULTFLCKAIQNTST-----LHLQSLCLFLAHLFLVAVI--DQTG 115
QY 334 SDAACWARGAVHYFLICAPTWMGLEAFHYLYLA--VRVNTYFGHYFLKSL--VGWGL 389
Db 116 HVLCSIIAGTLHYLYLATFTWMLLEALYLFRTARNLTVVNYSSINRPMKLMFPVGV 175
QY 390 PALMVIGTGSANSYGLYTRDRENTSLELCWFR-EGTMYALYTVHGYFLITFLGMV 448
Db 176 PA-VTVAISAASRPHLYTPSR-----CWLQPEKGFINGFLGPVCAIFSANLVFLV 226
QY 449 VVALVWVKIFTLSRATYKRGKNNKVKVLTILGLSSLGVGTWGLAIFT--PLGLSTVYIF 506
Db 227 TLWLKLNLSLN--SEVSTLRNTRMLAFKATAQLFILGCTWCLGLILQVGAARVMAYLF 284
QY 507 ALFNSLOGVFICCWFTIYLPQS 529
Db 285 TIINSLOGVFI--FLVYCLLSQ 304

RESULT 5
US-08-852-806-2
; Sequence 2, Application US/08852806
; Patent No. 5874245
; GENERAL INFORMATION:
; APPLICANT: Shoji Fukusumi
; APPLICANT: Shuji Hinuma
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,806
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017,915
; FILING DATE: 16 MAY 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: TAK50002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-852-806-2

Query Match 10.2%; Score 292.5; DB 2; Length 1052;
Best Local Similarity 28.2%; Pred. No. 1.1e-20;
Matches 127; Conservative 57; Mismatches 186; Indels 81; Gaps 23;

QY 118 RQVMKBDK-PDRVRLPKS-LFRSLPGRNSVVRVLAFTI-LDIGPGLFKGP-----RLG 169
Db 245 RRGKDWVRSEDRLELPKEVLSLSPGKPTSGAGSGRGCGTVPVPGHSHQRL 304
QY 170 LGDGS-----GVLNRLVGLSVGMHVTKLAEPLEIVFSHORPP-----PNMFLT-- 214
Db 305 PADPDESSYEVIGAVLYRTGLILPPR-PPLAVTSRVMTVTVPPTQPAEPLIIVLS 363
QY 215 -----CVFWDVTK--GTGDNSSGECST-EVRPEGTVCCHLTFALLRPTLD 261
Db 364 YIINGTDPHCASWYSRADASSDQDWTENCQPLETAAHTRCQCQHLSTFAVLAQPPK 423
QY 262 QSTVHILTRIS---QAGCGVSMIFLAFTIILY-AFURLSRERKSDAPKIHVALGSLF 317
Db 424 -LTLEAGSPVPLVIGCAVSCMALITLAIYAAFWRF-----IKSE---RSILNFCULS 475
QY 318 LNLAFVLNVGSGKSDAACWARGAVHYFLICAPTWMGLEAFHYLYLA--VRVNTYFG 376

Db 476 ILASNILILVQSRVLSKGVCTMTAAFLHFFFLSFCHVLTARQSVYLAIVIGRMTRILVR 535
QY 377 HYFLKLSLVGGLPALMV---IGTGSANSYGLYTIIRDRENTSLELCWFR-EGTMYALY 432
Db 536 KRFLCL--GWGLPALVAVSVGFTRTKYG-----TSSYCWLSLEGGLLYAFV 581
QY 433 ITVHGYYLITFLFGMVVLAVVWKIFTLRSATAVKRGKNEKKVYLTLG---LSSLVGV 489
Db 582 GPAAVIVLNNMLIGIIVFN-----KLMARDGISDKSKKORAGASLWSSCVLPLLA 634
QY 490 WGLAIFTPLGLSTVY---IFALFNSLOGVFI 517
Db 635 WMSAVLMTDRRSVLQALFAVNSAQGFVI 665

RESULT 6
US-09-163-669-2
; Sequence 2, Application US/09163669
; Patent No. 6111076
; GENERAL INFORMATION:
; APPLICANT: FUKUSUMI, SHOJI
; APPLICANT: HINUMA, SHUJI
; APPLICANT: FUJII, RYO
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR (HIBCD07)
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/163,669
; FILING DATE: 30-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/852,806
; FILING DATE: 07-MAY-1997
; APPLICATION NUMBER: 60/017,915
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: TAK-50002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-163-669-2

Query Match 10.2%; Score 292.5; DB 3; Length 1052;
Best Local Similarity 28.2%; Pred. No. 1.1e-20;
Matches 127; Conservative 57; Mismatches 186; Indels 81; Gaps 23;

QY 118 ROYMKDEK-PPDRVRLPKS-LFRLSLPGNRSVVRLAVTI-LDIGPGTFLFKGP-----RLG 169
Db 245 RRGKMDWRHSEDRLFLPKVELSLSSPGKPATSGAUGSPGRGPGTIVPPGPGHSHORLL 304

QY 170 LGDGS-----GVLANRLVGLSVGQMHVTKLAEPLEIVFSHORPB-----PNMTLT-- 214
Db 305 PADDESSYFVIGAVLYRTGLLPPPR-PPLAVTSRVMTVTRPTQPPAEPPLIIVELS 363
QY 215 -----CVFWDVTK--GTTGDMSSGECST-EVRPEGTVCCDHLFFALLRLPTLD 261
Db 364 YIINGTTDPHCASWDYSRADASSGDWDTCQTLFQAATRCQCOHLSSTFAVLAOPPKD 423
QY 262 QSTVHLITRIS---QAGCGVSMIFLAFTIILY-AFURLSRERFKSDEAPKIHVALGGSLF 317
Db 424 -LTLELAGSPVPLVICCAVSCMALLTLAIYAAWFR-----IKSE---RSIILLNFCLS 475
QY 318 LLNLAEVLNVGSGKSDAACWARGAVHYFLLCFTWGLAEAFHLYLLAV-RVNTYFG 376
Db 476 ILASNILILVQSRVLSKGVCTMTAAFLHFFFLSFCHVLTARQSVYLAIVIGRMTRILVR 535
QY 377 HYFLKLSLVGGLPALMV---IGTGSANSYGLYTIIRDRENTSLELCWFR-EGTMYALY 432
Db 536 KRFLCL--GWGLPALVAVSVGFTRTKYG-----TSSYCWLSLEGGLLYAFV 581
QY 433 ITVHGYYLITFLFGMVVLAVVWKIFTLRSATAVKRGKNEKKVYLTLG---LSSLVGV 489
Db 582 GPAAVIVLNNMLIGIIVFN-----KLMARDGISDKSKKORAGASLWSSCVLPLLA 634
QY 490 WGLAIFTPLGLSTVY---IFALFNSLOGVFI 517
Db 635 WMSAVLMTDRRSVLQALFAVNSAQGFVI 665

RESULT 7
US-09-370-098-5
; Sequence 5, Application US/09370098
; Patent No. 6303768
; GENERAL INFORMATION:
; APPLICANT: Lin, Yi-Jyun
; APPLICANT: Benzer, Seymour
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: METHUSELAH GENE, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 06618/343001
; CURRENT APPLICATION NUMBER: US/09/370,098
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 60/095,826
; PRIOR FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-370-098-5

Query Match 9.2%; Score 264; DB 4; Length 240;
Best Local Similarity 31.5%; Pred. No. 8.7e-19;
Matches 84; Conservative 37; Mismatches 98; Indels 48; Gaps 11;

QY 267 ILTRISQAGCGVSMIFLAFTIILYAFRLSRERFKSDEAPKIHVALGGSLFLNLAFV 326
Db 4 LLSVITWVGIVISLVCLAICTFCFL-----RGLQTDRTNTHKNLCINLFAELFL-- 56
QY 327 VGSKSGSDAACWARGAVHYFLLCFTWGLAEAFHLYLLAVRVNTYFG---HYFLKLS 383
Db 57 VGIDKTQYEVACPIFAGLLHYFFLAFAFSLCGLGVHLYLLVEVFESEYSRKYYILG-- 114
QY 384 LVWGGLPALMVIGTGSANSYGLYTIIRDRENTSLELCWFRGTTMYALYIVHGFILTF 443
Db 115 --GYCFPAL-VWGAIAAIDYRSYGT-----KACWLR-----VDNYFINSF 152
QY 444 LFGMVVLAVVWKIF-----TLSRATAVKRGKNE-----KKVLTLLGLSLVGV 490
Db 153 I-GPVSEVIVVNLVFLMVLHKMIRSSSVLKPDDSRDLNKSALGALIAILLGLGTWAF 211

QY 394 VGTGSANSYGLYITRDRENTSLELCWFREGTMYALYITVHGXYFLITFLGMMVLALV 453
Db 219 L-----GLAVGLDPEGYNPDFCW-----ISVHEPLIWSFA-GPVVLVIV 257
QY 454 V-WKIFTLSRATVAVKRGKRNKVVLTLLGLSS-----LVGVTW--GLAFTPLGLSTVYI 505
Db 258 MNGTWFLLAARTSCSTGQREAKTSALTIRSFLLLLLVLSASWLFGLLAVNHSILAFHYL 317
QY 506 FALFNSLQGVFICCFWFTIYLPQSQTIVSSSTAR 539
Db 318 HAGLGLQGLAVLLFCVLNADARAAMPACIGR 351
RESULT 10
US-08-982-412-2
; Sequence 2, Application US/08982412
; Patent No. 5958729
; GENERAL INFORMATION:
; APPLICANT: SOPPET, DANIEL R
; APPLICANT: LI, YI
; APPLICANT: ROSEN, CRAIG A
; APPLICANT: RUBEN, STEVEN M
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE,
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,412
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF161PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-982-412-2

Query Match 7.4%; Score 213; DB 2; Length 884;
Best Local Similarity 24.4%; Pred. No. 1e-12;
Matches 96; Conservative 61; Mismatches 163; Indels 74; Gaps 19;
QY 177 LNNRLVGLSV--GQMHVTKLAE-PLEIVFSHQRPPTMTLCVFDVTKGTT---GDWSS 230
Db 1 MNSPVSVAVFHGRNFRGLGILESPISLEFRLLOTANRKAICQWQD--PPGLAEQHGHWTA 59
QY 231 EGCSTVRPEGT--VCCCDHLTFALL-----RPTLDQST-----VHILTRISQACGV 278
Db 60 RDCEL-VHRNGSHARCRCSTGTGTGVLMDASPRERLEGLELLAVFTHVVAVSVA--- 114
QY 279 SMIFLAFTIYAFRLRSRERFKSEDAPIHVALGSLFLNLAFLVNVGSGKSDAAC 338
Db 115 -----ALVLTARILLSLRLKS-NVRGIHANVAALGVAEALLFL--LGIHRTNQLVC 164
QY 339 WARGAVFHYLLCAFTWMGLEAFHYLLAVRVNT-----YFGHYELKLSLVGWGLPALM 393

Db 165 TAVAILLHYFFLSTFAWLFVQGLHLYRMQVEPRNVDRGAMRFYH-----ALGWGPVAVL 218
QY 394 VGTGSANSYGLYITRDRENTSLELCWFREGTMYALYITVHGXYFLITFLGMMVLALV 453
Db 219 L-----GLAVGLDPEGYNPDFCW-----ISVHEPLIWSFA-GPVVLVIV 257
QY 454 V-WKIFTLSRATVAVKRGKRNKVVLTLLGLSS-----LVGVTW--GLAFTPLGLSTVYI 505
Db 258 MNGTWFLLAARTSCSTGQREAKTSALTIRSFLLLLLVLSASWLFGLLAVNHSILAFHYL 317
QY 506 FALFNSLQGVFICCFWFTIYLPQSQTIVSSSTAR 539
Db 318 HAGLGLQGLAVLLFCVLNADARAAMPACIGR 351
RESULT 11
US-09-370-098-6
; Sequence 6, Application US/09370098
; Patent No. 6303768
; GENERAL INFORMATION:
; APPLICANT: Benzer, Seymour
; APPLICANT: Lin, Yi-Jyun
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: METHUSELAH GENE, COMPOSITIONS AND
; FILE REFERENCE: 06618/343001
; CURRENT APPLICATION NUMBER: US/09/370,098
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 60/095,826
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-370-098-6

Query Match 6.9%; Score 199; DB 4; Length 235;
Best Local Similarity 27.0%; Pred. No. 3.6e-12;
Matches 71; Conservative 41; Mismatches 93; Indels 58; Gaps 14;
QY 278 VSMIFLAFTIYAFRLRSRERFKSEDAPIHVALGSLFLNLAFLVNVGSGKSDAA 337
Db 2 ISLVCLALAIATFLLCRAVQNH-----NTYMHLCVCLFLAKILFL--TGDKTDNQT 54
QY 338 CWARGAVFHYLLCAFTWMGLEAFHYLLA--VRVNTFYGHYFLK---LSLVGWGLPALM 393
Db 55 CAIIAGFLHYFLACFFWMLVEAVMLFMVRNLKVYN-YFSSRNKMLHLCAFGYGLPVLV 113
QY 394 VGTGSA--NSYGLYTIRDRENTSLELCWFREGTMYALYITVHGXYFLITFLGMM 447
Db 114 VIISASVQPRGYGMH-----NR-----CWLNTETG-----FIWSFLGPVCMII 151
QY 448 -VYLAIVVMKIFTL-----SRATVAVKRGKRNKVVLTLLGLSS--LVGVTWGLAIFT 496
Db 152 TMSVLLAWTLWLRKLCSSVSSEVSKLD-----TRLLTFKAIQAIIFGLCSWVLGIFQ 206
QY 497 --PLGLSTVYIFALFNSLQGVFI 517
Db 207 IGPLASIMAYLEFTIINSLOQAFI 229

RESULT 12
US-09-370-098-4
; Sequence 4, Application US/09370098
; Patent No. 6303768
; GENERAL INFORMATION:
; APPLICANT: Benzer, Seymour
; APPLICANT: Lin, Yi-Jyun
; APPLICANT: California Institute of Technology

; TITLE OF INVENTION: METHUSELAH GENE, COMPOSITIONS AND
; FILE REFERENCE: METHODS OF USE
; CURRENT APPLICATION NUMBER: 06618/343001
; CURRENT FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US/09/370,098
; PRIOR FILING DATE: 1999-08-06
; PRIOR FILING DATE: 60/095,826
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homosapiens
US-09-370-098-4

Query Match 6.9%; Score 198.5; DB 4; Length 231;
Best Local Similarity 30.8%; Pred. No. 3.9e-12;
Matches 68; Conservative 26; Mismatches 96; Indels 31; Gaps 11;
QY 308 IHVALGSLFLNLAFLVNVSGSGSDAACWARGAVFHYFLLCATWGMGLEAFHLYLLA 367
Db 25 IHLHLCICLFVGSTIFLAGTENEGGVGLRCLRVAGLLHYCGLAACWMSLEGLYFLV 84
QY 368 VRVENTYFHYFLKLSLVGWLGPALMVIGTGA---NSYGLYTRDRENTSLELCW--F 422
Db 85 VRVFOGQ-GLSTRWLCIGVGP-LLIVGVSAIYSKGYG---RPR-----YCWLD 131
QY 423 REGTMTYALYTHVGYFLIFLFGMVVLAVVWKIFTLRSRATAVKRGKNNKK--VLTLL 480
Db 132 EQGFLWSFL-----GPVFTILCNNAVIFVTYVWKL--TKFSEINPDMMKKKARALTIT 184
QY 481 GLSS--LVGVTW--GLAIFTPLGLSTVYIFALFNSLQGVPI 517
Db 185 AIAQLFLGCTWVFGLEIFDRSLRLVLYVFTILNCLQGAFL 225

RESULT 13
US-08-110-286A-2
; Sequence 2; Application US/08110286A
; Patent No. 5728545
; GENERAL INFORMATION:
; APPLICANT: Perrin, Marilyn H.
; APPLICANT: Chen, Ruoping
; APPLICANT: Lewis, Kathy A.
; APPLICANT: Vale Jr., Wylie W.
; APPLICANT: Donaldson, Cynthia J.
; TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,286A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US/08/079,320
; FILING DATE: 18-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9439

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-286A-2

Query Match 6.2%; Score 177.5; DB 1; Length 415;
Best Local Similarity 26.2%; Pred. No. 1.3e-09;
Matches 73; Conservative 40; Mismatches 117; Indels 49; Gaps 12;
QY 266 HILTRISQAGCGVSMIFLAFTIILYAFRLSRERPKSEDAKIHVALGSLFLLMLAFV 325
Db 117 HVAVIINYLGHICISLVALLVAFVL--FLRLSIRCLRN---IIHWNLSAFILRNATWV 171
QY 326 -----NVSGSGSGSDAACWARG--GAVFHYFLLCATWGMGLEAFHLYLLAVRV 371
Db 172 VOLTMSPEVHQSNVC-----WCRLVTAAYNFVHTNFFWFGGCGYLHTAIVLY 221
QY 372 NTYFGHYFLKLSLVGWLGPALMVIGTGSANSYGLYTRDRENTSLELCWFREGTMYAL 431
Db 222 STDRLRKWMFI-CICGWVPFPIIWAIGKLY-----YDN-----EKCWFGKRGVYTD 269
QY 432 YTVHGYFLITLFGMVVLAVVWKIFTLRSRATAVKRGKNNKKVLTLLGLSSLVGVTVG 491
Db 270 YI-YOGPMILVLLINFIPLFNIVRILMTKLRASTTSETIYQYKAVKATVLLPLLGITYM 328
QY 492 LAIFTPLGLSTV--YIFALFN---SLOGVFICCWFTIL 524
Db 329 LFFVNP-GEDEVSRVVFYFNSEFLESGFQFFVSVFYCF 366

RESULT 14
US-08-981-189B-11
; Sequence 11; Application US/08981189B
; Patent No. 6214797
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UROCORTIN PEPTIDES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITCH, EVEN, TABIN & FLANNERY
; STREET: 120 S. LaSalle Street, Suite 1600
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,189B
; FILING DATE: 10-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,144
; FILING DATE: 13-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,223
; FILING DATE: 11-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Schumann, James J.
; REGISTRATION NUMBER: 20,856
; REFERENCE/DOCKET NUMBER: 57611
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-552-1311
; TELEFAX: 858-552-0095

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..431
OTHER INFORMATION: /note= "Product-mouse heart derived
PUBLICATION INFORMATION:
AUTHORS: Perrin, Marilyn
AUTHORS: Donaldson, Cynthia
AUTHORS: Chen, Ruoping
AUTHORS: Blount, Amy
AUTHORS: Berggren, Travis
AUTHORS: Bileziklian, Louise
AUTHORS: Sawchenko, Paul
AUTHORS: Vale, Wylie
TITLE: Identification of a second
TITLE: corticotropin-releasing factor receptor gene and
TITLE: characterization of a cDNA expressed in heart
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 92
PAGES: 2969-2973
DATE: March-1995
US-08-981-189B-11

Query Match 6.28; Score 177.5; DB 4; Length 431;
Best Local Similarity 27.88; Pred. No. 1.4e-09;
Matches 73; Conservative 41; Mismatches 118; Indels 31; Gaps 11;
QY 271 ISQAGCGVSMIFLAFTIILYAFRLSRERFKSEDAPKIHVALGSGFLNLA-FLVNV-- 327
DB 139 VNYLGHCVSVVALVAFLFLVLRISRL-----RNVIHNLITFTILRNIAFWLQLID 193
QY 328 GSGKSGDAACWARGAVFHYFLLCFTWGLAFHLYLLAVRVNTYFGHYFLKLSLVGW 387
DB 194 HEVHEGNEVWCRCITTFINFYVVTFFWFMFVEGCVLHTAIVMTYSTEHLRKWFL-FIGW 252
QY 388 GLPALMVIGTGSANSYGLYTIRDRENRTSLCWFREGTMYALYIVHGYFLITFLFGM 447
DB 253 CIPCPITIAWAVGKLY-----YEN-----EQCFWKEAGDLVDYI-YQGPVMLVLLNF 300
QY 448 VVALVVMKIFTLRSATAVKRGKRNKKVLTLLGLSSLVGVTWGLAIFTP--LGLSTVYI 505
DB 301 VFLENIVRILMTKLRASTTSETIOYRKAVKATLVLLPLLGITYMLFFVNPGEDDLQI-V 359
QY 506 FALFN---SLOGVFIC---CWF 521
DB 360 FIYFNSFLQSFQGFVSVFYCF 382

RESULT 15
US-08-381-433A-2
Sequence 2, Application US/08381433A
Patent No. 5786203
GENERAL INFORMATION:
APPLICANT: Lovenberg, Timothy W.
APPLICANT: Oltersdorff, Tilman
APPLICANT: Liaw, Chen
APPLICANT: Grigorliadis, Dimitri E.
APPLICANT: DeSouza, Errol B.
TITLE OF INVENTION: CORTICOTROPIN RELEASING FACTOR 2
TITLE OF INVENTION: RECEPTORS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/381/433A
FILING DATE: 31-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 690068.401C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-381-433A-2

Query Match 6.18; Score 174.5; DB 1; Length 431;
Best Local Similarity 27.48; Pred. No. 2.9e-09;
Matches 72; Conservative 41; Mismatches 119; Indels 31; Gaps 11;
QY 271 ISQAGCGVSMIFLAFTIILYAFRLSRERFKSEDAPKIHVALGSGFLNLA-FLVNV-- 327
DB 139 INYLGHCVSVVALVAFLFLVLRISRL-----RNVIHNLITFTILRNITWFLQLID 193
QY 328 GSGKSGDAACWARGAVFHYFLLCFTWGLAFHLYLLAVRVNTYFGHYFLKLSLVGW 387
DB 194 HEVHEGNEVWCRCVTTIFNYFVVTFFWFMFVEGCVLHTAIVMTYSTEHLRKWFL-FIGW 252
QY 388 GLPALMVIGTGSANSYGLYTIRDRENRTSLCWFREGTMYALYIVHGYFLITFLFGM 447
DB 253 CIPCPITIAWAVGKLY-----YEN-----EQCFWKEAGDLVDYI-YQGPVMLVLLNF 300
QY 448 VVALVVMKIFTLRSATAVKRGKRNKKVLTLLGLSSLVGVTWGLAIFTP--LGLSTVYI 505
DB 301 VFLENIVRILMTKLRASTTSETIOYRKAVKATLVLLPLLGITYMLFFVNPGEDDLQI-V 359
QY 506 FALFN---SLOGVFIC---CWF 521
DB 360 FIYFNSFLQSFQGFVSVFYCF 382

Search completed: June 11, 2002, 22:25:18
Job time: 3235 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: June 11, 2002, 21:27:38 ; Search time 44.13 Seconds
(without alignments)
1195.401 Million cell updates/sec

Title: us-09-733-387-44
Perfect score: 2878
Sequence: 1 MATPRGLGALLLLPTSG.....STTVSSSTARLDQAHSQEQ 549
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	383.5	13.3	1230	2	CL3AB protein - ra
2	383.5	13.3	1231	2	latrophilin-3, spl
3	383.5	13.3	1273	2	CL3AC protein - ra
4	383.5	13.3	1274	2	latrophilin-3, spl
5	383.5	13.3	1298	2	CL3BB protein - ra
6	383.5	13.3	1299	2	latrophilin-3, spl
7	383.5	13.3	1341	2	CL3BC protein - ra
8	383.5	13.3	1342	2	latrophilin-3, spl
9	383.5	13.3	1459	2	CL3AA protein - ra
10	383.5	13.3	1503	2	latrophilin-3, spl
11	383.5	13.3	1527	2	CL3BA protein - ra
12	383.5	13.3	1571	2	latrophilin-3, spl
13	379	13.2	1240	2	latrophilin-3, spl
14	379	13.2	1283	2	latrophilin-3, spl
15	379	13.2	1308	2	latrophilin-3, spl
16	379	13.2	1351	2	latrophilin-3, spl
17	379	13.2	1512	2	latrophilin-3, spl
18	379	13.2	1550	2	latrophilin-3, spl
19	379	13.2	1580	2	latrophilin-3, spl
20	370.5	12.9	1467	2	alpha-latotoxin r
21	370.5	12.9	1472	2	latrophilin-1, bra
22	367	12.8	1420	2	latrophilin-1, bra
23	367	12.8	1452	2	CL2AB protein - ra
24	367	12.8	1463	2	CL2AA protein - ra
25	365.5	12.7	1466	2	CL2AC protein - ra
26	365.5	12.7	1471	2	CL1AA protein - ra
27	365.5	12.7	1510	2	CL1BA protein - ra
28	365.5	12.7	1515	2	CL1AB protein - ra
29	359.5	12.5	1354	2	latrophilin-2 (spl)

30 359.5 12.5 1397 2 T18377 latrophilin-2 (spl)
31 359.5 12.5 1420 2 T18385 latrophilin-2 (spl)
32 359.5 12.5 1463 2 T18386 latrophilin-2 (spl)
33 355.5 12.4 1435 2 T46611 CL2BA protein - ra
34 355.5 12.4 1467 2 T17160 CL2BA protein - ra
35 355.5 12.4 1478 2 T17185 CL2BC protein - ra
36 353 12.3 1341 2 T18301 latrophilin-2, spl
37 353 12.3 1384 2 T18366 latrophilin-2, spl
38 353 12.3 1407 2 T18381 latrophilin-2 (spl)
39 353 12.3 1450 2 T18382 latrophilin-2 (spl)
40 351 12.2 1487 2 T14324 alpha-latotoxin r
41 348 12.1 1369 2 T18379 latrophilin-2 (spl)
42 348 12.1 1412 2 T18380 latrophilin-2 (spl)
43 348 12.1 1435 2 T18387 latrophilin-2 (spl)
44 348 12.1 1478 2 T18388 latrophilin-2 (spl)
45 341.5 11.9 1356 2 T18367 latrophilin-2, spl

RESULT 1
T17187
CL3AB protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T17187
R:Sugita, S.; Ichchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: Z18712
A:Accession: T17187
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1230 <SUG>
A:Cross-references: EMBL:AF081155; NID:g3695136; PID:g3695137; PIDN:AAC62661.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent

Query Match	13.3%	Score	383.5;	DB	2;	Length	1230;
Best Local Similarity	27.6%	Pred. No.	2.5e-23;				
Matches	135;	Conservative	77;	Mismatches	182;	Indels	95;
Gaps	25;						
QY	77	HLMKGLTQKVTNPFKALVQNLST-NTAEDFYFSLPS----	QVPRQVKDEKPPDR	130			
Db	682	NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDKFPENTGHSTIQLSANTLKQN-----		734			
QY	131	VRLPKSLFSLPGRSVVRLAVTIL-DIGPGTLFKGPRGLG-----	DGSGVLNNRLVG	183			
Db	735	-----GRNGEIRVAFVLYNNGPYLSTENASKLGTAMSTNHSVIVNSPVIT	782				
QY	184	LSVGOMHVTK--LAEPLEIVFSH-QRPPNMTLTCVFDVDTKGT-TGWSSEGS-TEVR	238				
Db	783	AAINTEFNSKIVLADVPVTVTKHKIQSEENFNPCNSFWSKRTMTGYSTGCRLLTTN	842				
QY	239	PBGTVCCDHLTFPALLLR--PTLDQSTVH--ILTRISQACGGVSMIFLATIILYAFLR	294				
Db	843	KHTTCSNHLTNFAVLMAHVEVKHSDAHDLLDVTWVGILLSVCLLICITFFCFF--	901				
QY	295	LSRERFKSEDAPIKHVALGSLFLNLAFVNVGSGKSDAACWARGAVFHYELLCRAFT	354				
Db	902	----RGLQSDRNTIHNKLCISLVAELLFL--IGINRTDQPIACAVFAALLHFFFLAFT	955				
QY	355	WMGLEAFHLYLLAVRVFTYFGH--YFLKLSLVGWGLPALMVIQTGSAN--SYGLYTIRD	410				
Db	956	WNFLEGVOLYIMLVFEFSEHSRRKIFY--LVGYGMPALIVAVSAADYRSYG-----	1006				
QY	411	RENRTSLCLNFRGRTTWALYITVHGVELTTF-----LFGMVVLALVWVKIFLSR	462				
Db	1007	-----TDKVCWRLDT-----YFWSFIDPATLIIMLVNIFGLALYKMF---H	1047				
QY	463	ATAV--KERG--KNRKK-VLTLLGLSLVGVTV--GLAIFTPLGLSTVYIFALFNSLQGV	515				

```

Db      1048  HTAILPESGCLDNIKSWVIGAIALCLGLTWAFGLMYINESTVIMAYLFTIFNSIQGM 1107
QY      516  FICCWFTIL 524
        || : :|
Db      1108  FIFIFHCVL 1116

```

RESULT 2

T18390

latrophilin-3, splice variant abag, brain-specific - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T18390

R:Hatsushita, H.; Lelianova, V.G.; Ushkaryov, Y.A.

FEB5 Lett. 443, 348-352, 1999

A:Title: The latrophilin family: multiply spliced G protein-coupled receptors

A:Reference number: Z18869; MUID:99149828

A:Accession: T18390

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1331 <MAT>

A:Cross-references: EMBL:AF111086; NID:g4164054; PID:g4164055; PIDN:AAD0532

C:Superfamily: alpha-latrototoxin receptor, calcium-independent

C:Keywords: alternative splicing; G protein-coupled receptor

[illegible]

```
Qy    463 ATAV--KERG--KNRKK-VLTLLGLSSLVGVTVW--GLAIFTPLGLSTVYVFALFNSLQGV 515
      || : - | : : | :| :|| ||   : || :||||| : 
Db    1049 HTAILKPESCLDNIKSWIGAIALLCILGLTFAFGLMVINESTVMAYLFITFNLSQQM 1108
      || : - | : : | :| :|| ||   : || :||||| : 

Qy    516 FICCWFTIL 524
      || : : | 
Db    1109 FIPIFHCVL 1117
      || : : | 

RESULT          3
TL17188
CL3AC protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: TL17188
```

R; Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C. submitted to the EMBL Data Library, July 1998
A; Description: CL family.
A; Reference number: Z18712
A; Accession: T17188
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: mRNA
A; Residue: 1-1273 <SUG>
A; Cross-references: EMBL:AF081156; NID:g3695138; PID:g3695139; PIDN:AAC62662.1
C; Superfamily: alpha-latrototoxin receptor, calcium-independent

[illegible]

RESULT 4
T18391
latrophilin-3, splice variant abah, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18391
R:Matsumura, H.; Lelanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: Z18869; MUID:99148828
A:Accession: T18391
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1274 <MAT>
A:Cross-references: EMBL:AF111087; NID:q4164056; PID:q4164057; PIDN:AA05323.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

Query Match
13.3%; Score 383.5; DB 2; Length 1274;

Best Local Similarity 27.6%; Pred. No. 2.7e-23;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

```

QY 77 HLMKEGLTKVNTPELTKALVQNLST-NTAEDFYFSLPS-----QVPRQVMKDEKPPDR 130
Db 683 NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDKFPENTGHGSTIQLSANTLKQ-----735
QY 131 VRLPKSLFRSLPGRNSVVRVLAUTIL-DIGPGTLFKPRGLG-----DGSGLNNRLVG 183
Db 736 -----GRGEIRVAVFLYNNLGPYLSTENASMKLGTEAMSTNHSVIVNSPVIT 783
QY 184 LSVGQMHVTK--LAEPLEIVFSH-QRPPNNMTLTCVFDVTKGT-TGDMSSGCS-TEVR 238
Db 784 AAINKEFSNKVYLADPVVFTVKHIKQSEENFNPCNSFWSYKRTWTGYWSTQGCRLTTN 843
QY 239 PEGTVCCDHLTFALLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
Db 844 KTHITCSNHLTNFAVLMAHVEVKHSDAHDLLDVLTVWGLLSLVCLLICIFTCFF-902
QY 295 LSRERFKSEDPKIHVALGSLFLNLNLAFLVNVGSGSKSDAACWARGAVFHYFLCAFT 354
Db 903 -----RGLQSDRNTIHKNLICISLFAELLFL--IGINRTDQPIACAVFAALLHFFFLAAFT 956
QY 355 WMGLEAFHLYLLAVRVNTFYGH--YFLKLSLVGWLGPALMVIGTGSAN--SYGLYTIRD 410
Db 957 WMFLEGVQLYIMLVEVFESHSRRKYFY---LVGYGMPALIVAVSAADVRSYG-----1007
QY 411 RENRTSLELCWREGTMYALYITVHGYPFLITF-----LFGMVVLALVVMKIFTLR 462
Db 1008 -----TDKVCWLRLDT-----YFWSFPGPATLIIMNLVIFGLIALYKMF---H 1048
QY 463 ATAV--KERG--KNRKK-VLTLLGSLVGVTV--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1049 HTAILKPESGCLDNKISWIGAIALLCLGLTWAFGLMYNESTVIMAYLFTFNSLQGM 1108
QY 516 FICCWFTIL 524
Db 1109 FIFIFHCVL 1117

RESULT 5
T17199
CL3BB protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T17199
R:Sugita, S.; Ichchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: Z18712
A:Accession: T17199
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-1298 <SUG>
A:Cross-references: EMBL:AF081158; NID:g3695142; PID:g3695143; PIDN:AACG2664.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent

```

```

Query Match 13.3%; Score 383.5; DB 2; Length 1298;
Best Local Similarity 27.6%; Pred. No. 2.7e-23;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

QY 77 HLMKEGLTKVNTPELTKALVQNLST-NTAEDFYFSLPS-----QVPRQVMKDEKPPDR 130
Db 750 NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDKFPENTGHGSTIQLSANTLKQ-----802
QY 131 VRLPKSLFRSLPGRNSVVRVLAUTIL-DIGPGTLFKPRGLG-----DGSGLNNRLVG 183
Db 803 -----GRGEIRVAVFLYNNLGPYLSTENASMKLGTEAMSTNHSVIVNSPVIT 850
QY 184 LSVGQMHVTK--LAEPLEIVFSH-QRPPNNMTLTCVFDVTKGT-TGDMSSGCS-TEVR 238
Db 851 AAINKEFSNKVYLADPVVFTVKHIKQSEENFNPCNSFWSYKRTWTGYWSTQGCRLTTN 910

```

```

QY 239 PEGTVCCDHLTFALLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
Db 911 KTHITCSNHLTNFAVLMAHVEVKHSDAHDLLDVLTVWGLLSLVCLLICIFTCFF-969
QY 295 LSRERFKSEDPKIHVALGSLFLNLNLAFLVNVGSGSKSDAACWARGAVFHYFLCAFT 354
Db 970 -----RGLQSDRNTIHKNLICISLFAELLFL--IGINRTDQPIACAVFAALLHFFFLAAFT 1023
QY 355 WMGLEAFHLYLLAVRVNTFYGH--YFLKLSLVGWLGPALMVIGTGSAN--SYGLYTIRD 410
Db 1024 WMFLEGVQLYIMLVEVFESHSRRKYFY---LVGYGMPALIVAVSAADVRSYG-----1074
QY 411 RENRTSLELCWREGTMYALYITVHGYPFLITF-----LFGMVVLALVVMKIFTLR 462
Db 1075 -----TDKVCWLRLDT-----YFWSFPGPATLIIMNLVIFGLIALYKMF---H 1115
QY 463 ATAV--KERG--KNRKK-VLTLLGSLVGVTV--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1116 HTAILKPESGCLDNKISWIGAIALLCLGLTWAFGLMYNESTVIMAYLFTFNSLQGM 1175
QY 516 FICCWFTIL 524
Db 1176 FIFIFHCVL 1184

```

```

RESULT 6
T18398
latrophilin-3, splice variant bbag, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18398
R:Matsumita, H.; Lelianaova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: Z18869; MUID:99148828
A:Accession: T18398
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-1299 <MAT>
A:Cross-references: EMBL:AF111092; NID:g4164066; PID:g4164067; PIDN:AAD05328.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

```

```

Query Match 13.3%; Score 383.5; DB 2; Length 1299;
Best Local Similarity 27.6%; Pred. No. 2.7e-23;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

```

```

QY 77 HLMKEGLTKVNTPELTKALVQNLST-NTAEDFYFSLPS-----QVPRQVMKDEKPPDR 130
Db 751 NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDKFPENTGHGSTIQLSANTLKQ-----803
QY 131 VRLPKSLFRSLPGRNSVVRVLAUTIL-DIGPGTLFKPRGLG-----DGSGLNNRLVG 183
Db 804 -----GRGEIRVAVFLYNNLGPYLSTENASMKLGTEAMSTNHSVIVNSPVIT 851
QY 184 LSVGQMHVTK--LAEPLEIVFSH-QRPPNNMTLTCVFDVTKGT-TGDMSSGCS-TEVR 238
Db 852 AAINKEFSNKVYLADPVVFTVKHIKQSEENFNPCNSFWSYKRTWTGYWSTQGCRLTTN 911
QY 239 PEGTVCCDHLTFALLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
Db 912 KTHITCSNHLTNFAVLMAHVEVKHSDAHDLLDVLTVWGLLSLVCLLICIFTCFF-970
QY 295 LSRERFKSEDPKIHVALGSLFLNLNLAFLVNVGSGSKSDAACWARGAVFHYFLCAFT 354
Db 971 -----RGLQSDRNTIHKNLICISLFAELLFL--IGINRTDQPIACAVFAALLHFFFLAAFT 1024
QY 355 WMGLEAFHLYLLAVRVNTFYGH--YFLKLSLVGWLGPALMVIGTGSAN--SYGLYTIRD 410
Db 1025 WMFLEGVQLYIMLVEVFESHSRRKYFY---LVGYGMPALIVAVSAADVRSYG-----1075

```

QY 411 RENRTSLELCWFREGTMYALYITVHGFLITF-----LFGMVVLALVVMKIFTLR 462
Db 1076 -----TDKVCWLRDT-----YFIWSFGPATLIIMLVIFGLIALYKMF---H 1116
QY 463 ATAV--KERG--KNRKK-VLTLLGLSSLVGVW--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1117 HTAILKPESGCLDNKSWVIGAIALLCLGLTWAFGLMYINESTVIMAYLFTIFNSLQGM 1176
QY 516 FICCWFTIL 524
Db 1177 FIFIFHCVL 1185
RESULT 7
CL3AA protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T17200
R:Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family,
A:Reference number: Z18712
A:Accession: T17200
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1341 <SUG>
A:Cross-references: EMBL:AF081159; NID:g3695144; PID:g3695145; PIDN:AAC62665.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
Query Match 13.3%; Score 383.5; DB 2; Length 1341;
Best Local Similarity 27.6%; Pred. No. 2.8e-23;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;
QY 77 HLMKEGLTQKVTNPFKALYQNLST-NTAEDFYFSLEPS-----QVPROVMKDEKPPDR 130
Db 750 NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDLAFPTNGHSTIQLSANTLQON----- 802
QY 131 VRLPKSLFRSLPGNRSVRVLAIVTL-DIGPGTLFKGPRGLG-----DGSGLVNNRLVG 183
Db 803 -----GRNGEIRVAFVLYNNLPGYLSLTENASMKLGTEAMSTNHSVIVNSPVIT 850
QY 184 LSVGMHVTK--LAEPLEIVFESH-ORPPNMTLTCVFDVTKGT-TGDRSSECCS-TEVR 238
Db 851 AAINKEFSNKVYLADPVVTVKHKOSEENFNPCNCSFWSYKRTMTGYWSTGCRLLTTN 910
QY 239 PEGTVCCDHLTFALLR--PTLDQSTVH--ILTRISOAGCGVSMIFLAFTIILYAFLR 294
Db 911 KTHITCSNHLTNFVLMHAEVKHSDAHDLLDVIWVGILLSLVCLLICIFTCFF- 969
QY 295 LSREPKSEDAPIHVALGSLFLNLAFVNVGSGKSDAACWARGAVHYFLCAFT 354
Db 970 -----RGLQSDRNTIHNKLCISLFAELLF--IGINRTDQPIACAVFAALLHFFFLAAFT 1023
QY 355 WMGLEAFHLYLLAVRVNTYFEGH--YFLKLSLVGHLGALPMVIGTGSAN--SYGLYTIRD 410
Db 1024 WMFLEGVQLYIMLVFESEHSRRKIFY---LVGMPALIVAVSAADYRSYG----- 1074
QY 411 RENRTSLELCWFREGTMYALYITVHGFLITF-----LFGMVVLALVVMKIFTLR 462
Db 1075 -----TDKVCWLRDT-----YFIWSFGPATLIIMLVIFGLIALYKMF---H 1115
QY 463 ATAV--KERG--KNRKK-VLTLLGLSSLVGVW--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1116 HTAILKPESGCLDNKSWVIGAIALLCLGLTWAFGLMYINESTVIMAYLFTIFNSLQGM 1175
QY 516 FICCWFTIL 524
Db 1176 FIFIFHCVL 1184
RESULT 8

T18405
latrophilin-3, splice variant bbah, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18405
R:Natsushita, H.; Lellanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: Z18869; MUID:99148828
A:Accession: T18405
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1342 <MAT>
A:Cross-references: EMBL:AF111093; NID:g4164068; PID:g4164069; PIDN:AAO05329.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
Query Match 13.3%; Score 383.5; DB 2; Length 1342;
Best Local Similarity 27.6%; Pred. No. 2.8e-23;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;
QY 77 HLMKEGLTQKVTNPFKALYQNLST-NTAEDFYFSLEPS-----QVPROVMKDEKPPDR 130
Db 751 NLLKTDIVRE-NTDNIQLEVARLSTEGNLEDLKFPTNGHSTIQLSANTLQON----- 803
QY 131 VRLPKSLFRSLPGNRSVRVLAIVTL-DIGPGTLFKGPRGLG-----DGSGLVNNRLVG 183
Db 804 -----GRNGEIRVAFVLYNNLPGYLSLTENASMKLGTEAMSTNHSVIVNSPVIT 851
QY 184 LSVGMHVTK--LAEPLEIVFESH-ORPPNMTLTCVFDVTKGT-TGDRSSECCS-TEVR 238
Db 852 AAINKEFSNKVYLADPVVTVKHKOSEENFNPCNCSFWSYKRTMTGYWSTGCRLLTTN 911
QY 239 PEGTVCCDHLTFALLR--PTLDQSTVH--ILTRISOAGCGVSMIFLAFTIILYAFLR 294
Db 912 KTHITCSNHLTNFVLMHAEVKHSDAHDLLDVIWVGILLSLVCLLICIFTCFF- 970
QY 295 LSREPKSEDAPIHVALGSLFLNLAFVNVGSGKSDAACWARGAVHYFLCAFT 354
Db 971 -----RGLQSDRNTIHNKLCISLFAELLF--IGINRTDQPIACAVFAALLHFFFLAAFT 1024
QY 355 WMGLEAFHLYLLAVRVNTYFEGH--YFLKLSLVGHLGALPMVIGTGSAN--SYGLYTIRD 410
Db 1025 WMFLEGVQLYIMLVFESEHSRRKIFY---LVGMPALIVAVSAADYRSYG----- 1075
QY 411 RENRTSLELCWFREGTMYALYITVHGFLITF-----LFGMVVLALVVMKIFTLR 462
Db 1076 -----TDKVCWLRDT-----YFIWSFGPATLIIMLVIFGLIALYKMF---H 1116
QY 463 ATAV--KERG--KNRKK-VLTLLGLSSLVGVW--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1117 HTAILKPESGCLDNKSWVIGAIALLCLGLTWAFGLMYINESTVIMAYLFTIFNSLQGM 1176
QY 516 FICCWFTIL 524
Db 1177 FIFIFHCVL 1185
RESULT 9
CL3AA protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T17186
R:Sugita, S.; Ichtchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family,
A:Reference number: Z18712
A:Accession: T17186
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1459 <SUG>
A:Cross-references: EMBL:AF081154; NID:g3695134; PID:g3695135; PIDN:AAC62660.1

[illegible]

1. **ИЗМЕНЕНИЕ ПОСРЕДСТВОМ КОТОРОГО** 2. **ВВЕДЕНИЯ ТЕХНОЛОГИИ ТИПОГРАФИИ** 3. **СС**

A:Reference number: Z18869; MUID:99148828

A:Accession: T18394

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1283 <MAT>

A:Cross-references: EMBL:AF111095; NID:g4164062; PID:g4164063; PIDN:AAD05326.1

C:Superfamily: alpha-latrotoxin receptor, calcium-independent

C:Keywords: alternative splicing; G protein-coupled receptor

Query Match 13.2%; Score 379; DB 2; Length 1283;

Best Local Similarity 26.5%; Pred. No. 6.3e-23;

Matches 131; Conservative 77; Mismatches 189; Indels 98; Gaps 22;

QY 77 HLMKEGLTKVNTPELKAALVONLST-NTAEDEVFSLEPS-----OVPRQVMKDEKPPDR 130

Db 683 NLKTDIVRE-NTDNIQLEVARLSTEGNLEDLKFPENTGHSTIQLSANTLKQN----- 735

QY 131 VRLPKSLFRSLPGNRVVRVLAFTIL-DIGPGTLFGKPRGLG-----DGSGLVNNRLVG 183

Db 736 -----GRNGEIRVAFVLYNNLGPYLSTENASMKLGTEAMSTNHSIVNSPVIT 783

QY 184 LSVQGMHVTK--LAEPLEIVFESH-QRPPNMTLTCVFDVTKGT-TGDWSSEGCs-TEVR 238

Db 784 AAINKEFSNKVYLADPWFTVKHIKQSEENFNPCSFWSYKRTMTGYWSTQCRLLTN 843

QY 239 PEGTVCCDHLTFALLLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294

Db 844 KTHHTCSNHLTNEAVLMAHVEVKHSDAHDLLDVTWVIGLLSLVCLLICITFCFF- 902

QY 295 LSREKFKSEDAPKIHVALGSLFLLNLAFLVNVGSGKSDAACWARGAVFHYFLICAPT 354

Db 903 ----RGLQSDRNTIHKNLCSLFAVELL--IGINRTDQPIACAVFAALLHFFFLAFT 956

QY 355 WMGLEAFHLYLLAVRVNTYFGH--YFLKLSLVGWLGPALMVIGTGSAN--SYGLYTIRD 410

Db 957 WMFEGVOLYIMLVEFESEHSRRKIFY---LVGIGMPALIVAVSAADVRSYG----- 1007

QY 411 RENRTSLELCWFREGTMYALYITVHGYYFLITF-----LFGMVVLAALVVMKIF---- 458

Db 1008 -----TDKVCWLRLDT-----YFIWSFIGPATLIIMLVIFLGIALYKMFHHTA 1051

QY 459 -----TLSRATAVKERGNKRVKLVTLGLSLVGVTVW--GLAIFTPLGLSTVYIFALF 509

Db 1052 ILKPESGCLDNINVEDNRPPIKSWVIGAIALLCLLGLTWAFGLMYINESTVIMAYLFTIF 1111

QY 510 NSLQGVFICCWFTIL 524

Db 1112 NSLQGMFIFIFHCVL 1126

RESULT 15

T18408

latrophilin-3, splice variant bbbg, brain-specific - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T18408

R:Matsushita, H.; Lelianaova, V.G.; Ushkaryov, Y.A.

FEBS Lett. 443, 348-352, 1999

A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with diffe

A:Reference number: Z18869; MUID:99148828

A:Accession: T18408

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1308 <MAT>

A:Cross-references: EMBL:AF111095; NID:g4164072; PID:g4164073; PIDN:AAD05331.1

C:Superfamily: alpha-latrotoxin receptor, calcium-independent

C:Keywords: alternative splicing; G protein-coupled receptor

Query Match 13.2%; Score 379; DB 2; Length 1308;

Best Local Similarity 26.5%; Pred. No. 6.5e-23;

Matches 131; Conservative 77; Mismatches 189; Indels 98; Gaps 22;

QY 77 HLMKEGLTKVNTPELKAALVONLST-NTAEDEVFSLEPS-----OVPRQVMKDEKPPDR 130

Db 751 NLKTDIVRE-NTDNIQLEVARLSTEGNLEDLKFPENTGHSTIQLSANTLKQN----- 803

QY 131 VRLPKSLFRSLPGNRVVRVLAFTIL-DIGPGTLFGKPRGLG-----DGSGLVNNRLVG 183

Db 804 -----GRNGEIRVAFVLYNNLGPYLSTENASMKLGTEAMSTNHSIVNSPVIT 851

QY 184 LSVQGMHVTK--LAEPLEIVFESH-QRPPNMTLTCVFDVTKGT-TGDWSSEGCs-TEVR 238

Db 852 AAINKEFSNKVYLADPWFTVKHIKQSEENFNPCSFWSYKRTMTGYWSTQCRLLTN 911

QY 239 PEGTVCCDHLTFALLLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294

Db 912 KTHHTCSNHLTNEAVLMAHVEVKHSDAHDLLDVTWVIGLLSLVCLLICITFCFF- 970

QY 295 LSREKFKSEDAPKIHVALGSLFLLNLAFLVNVGSGKSDAACWARGAVFHYFLICAPT 354

Db 971 ----RGLQSDRNTIHKNLCSLFAVELL--IGINRTDQPIACAVFAALLHFFFLAFT 1024

QY 355 WMGLEAFHLYLLAVRVNTYFGH--YFLKLSLVGWLGPALMVIGTGSAN--SYGLYTIRD 410

Db 1025 WMFEGVOLYIMLVEFESEHSRRKIFY---LVGIGMPALIVAVSAADVRSYG----- 1075

QY 411 RENRTSLELCWFREGTMYALYITVHGYYFLITF-----LFGMVVLAALVVMKIF---- 458

Db 1076 -----TDKVCWLRLDT-----YFIWSFIGPATLIIMLVIFLGIALYKMFHHTA 1119

QY 459 -----TLSRATAVKERGNKRVKLVTLGLSLVGVTVW--GLAIFTPLGLSTVYIFALF 509

Db 1120 ILKPESGCLDNINVEDNRPPIKSWVIGAIALLCLLGLTWAFGLMYINESTVIMAYLFTIF 1179

QY 510 NSLQGVFICCWFTIL 524

Db 1180 NSLQGMFIFIFHCVL 1194

Search completed: June 11, 2002, 22:24:34

Job time: 3416 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 22:23:31 ; Search time 25.32 seconds
(without alignments)
839.535 Million cell updates/sec

Title: US-09-733-387-44

Perfect score: 2878

Sequence: 1 MATPRGLGALLLLLLPTSG.....STTVSSSTARLDQAHSAQ 549

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	312	10.8	1584	1	BAIL_HUMAN
2	311	10.8	835	1	CD97_HUMAN
3	280.5	9.7	931	1	EMRL_MOUSE
4	277	9.6	1572	1	BAIL_HUMAN
5	270	9.4	1522	1	BAIL_HUMAN
6	254.5	8.8	886	1	EMRL_HUMAN
7	179.5	6.2	415	1	EMRL_SHEEP
8	177.5	6.2	431	1	CRF1_MOUSE
9	176.5	6.1	415	1	CRF1_MOUSE
10	173.5	6.0	411	1	CRF1_MOUSE
11	172.5	6.0	415	1	CRF1_MOUSE
12	168	5.8	441	1	DIHR_ACHDO
13	166	5.8	444	1	CRF1_HUMAN
14	165	5.7	415	1	CRF1_XENLA
15	161.5	5.6	413	1	CRF1_XENLA
16	161.5	5.6	420	1	CRF1_CHICK
17	161	5.6	517	1	MTH_DROXA
18	158	5.5	550	1	GLP2_RAT
19	157.5	5.5	515	1	CRF2_HUMAN
20	157	5.5	515	1	MTH_DROSI
21	156.5	5.4	458	1	MTH3_DROME
22	152.5	5.3	395	1	DIHR_MANSE
23	151	5.2	514	1	MTH_DROME
24	149.5	5.2	553	1	GLP2_HUMAN
25	149	5.2	468	1	MTH4_DROME
26	148.5	5.2	447	1	VIPR_CARAU
27	148.5	5.2	546	1	PTH2_RAT
28	148.5	5.2	585	1	PTH2_DIDMA
29	146.5	5.1	478	1	CAIR_CAVPO
30	142.5	5.0	550	1	PTH2_HUMAN
31	140.5	4.9	585	1	PTTR_PIG
32	140	4.9	676	1	MTH1_DROME
33	139	4.8	593	1	PTTR_HUMAN

34	137.5	4.8	437	1	VIPS_RAT
35	136.5	4.7	464	1	CGRR_RAT
36	135.5	4.7	489	1	GLPI_MOUSE
37	135	4.7	498	1	CALR_PIG
38	134	4.7	437	1	VIPS_MOUSE
39	134	4.7	474	1	CALR_RABIT
40	132.5	4.6	461	1	CGRR_HUMAN
41	132	4.6	466	1	GIPR_HUMAN
42	130.5	4.5	463	1	GLPI_HUMAN
43	130.5	4.5	468	1	PACR_HUMAN
44	130	4.5	591	1	PTTR_MOUSE
45	130	4.5	591	1	PTTR_RAT

ALIGNMENTS

RESULT 1
BAIL_HUMAN
ID BAIL_HUMAN STANDARD; PRT; 1584 AA.
AC Q14514;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 1 precursor.
GN BAIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98054121; PubMed=9393972;
RA Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,
RA Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.,
RT "A novel brain-specific p53-target gene, BAIL, containing
thrombospondin type 1 repeats inhibits experimental angiogenesis.";
RL Oncogene 15:2145-2150(1997).
RN [2]
RP INTERACTION WITH BAP1.
RX MEDLINE=98321173; PubMed=9647739;
RA Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,
RA Tokino T.;
RT "Cloning and characterization of BAI-associated protein 1: a PDZ
domain-containing protein that interacts with BAIL.";
RL Biochem. Biophys. Res. Commun. 247:597-604(1998).
CC -!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN
BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53
SIGNAL IN SUPPRESSION OF GLIOBLASTOMA. MAY FUNCTION IN CELL
ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.
CC -!- SUBUNIT: INTERACTS WITH BAP1.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE
CONCENTRATED AT CELL-CELL ADHESION SITES.
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED OR NO
EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES AND CANCER
TISSUES.
CC -!- INDUCTION: BY P53.
CC -!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT
CORNEA INDUCED BY BEGF.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 5 TSP TYPE-1 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005297; BAA23647.1; .
DR

DR MIM; 602682; --
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSPL.
DR InterPro; IPR001879; hormn_receptor.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp_1; 5.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; HormR; 1.
DR SMART; SM00209; TSPL; 5.
DR PROSITE; PS50221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECIP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECIP_F2_2; FALSE_NEG.
DR PROSITE; PS50227; G_PROTEIN_RECIP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECIP_F2_4; 1.
DR PROSITE; PS50092; TSPL; 5.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
Repeat; Cell adhesion.
KW SIGNAL
FT CHAIN 1 30
FT DOMAIN 31 1584
FT TRANSFAM 31 948
FT TRANSFAM 949 969
FT TRANSFAM 970 980
FT TRANSFAM 981 1001
FT TRANSFAM 1002 1008
FT TRANSFAM 1009 1029
FT TRANSFAM 1030 1052
FT TRANSFAM 1053 1073
FT TRANSFAM 1074 1093
FT TRANSFAM 1094 1114
FT TRANSFAM 1115 1136
FT TRANSFAM 1137 1157
FT TRANSFAM 1158 1166
FT TRANSFAM 1167 1187
FT TRANSFAM 1188 1584
FT TRANSFAM 261 316
FT TRANSFAM 354 408
FT TRANSFAM 409 463
FT TRANSFAM 467 521
FT TRANSFAM 522 576
FT TRANSFAM 581 938
FT TRANSFAM 1411 1422
FT TRANSFAM 1425 1430
FT SITE 231 233
FT DOMAIN 1365 1584
FT DOMAIN 1581 1584
FT CARBOHYD 64 64
FT CARBOHYD 401 401
FT CARBOHYD 607 607
FT CARBOHYD 692 692
FT CARBOHYD 844 844
FT CARBOHYD 877 877
FT CARBOHYD 881 881
FT SEQUENCE 1584 AA; 173531 MW; DEAF28C77874513 CRC64;
Query Match 10.8%; Score 312; DB 1; Length 1584;
Best Local Similarity 25.8%; Pred. No. 2.9e-15;
Matches 122; Conservative 70; Mismatches 180; Indels 100; Gaps 21;
Qy 96 VONLSTNTAEQFYSLEPSQVPMKDEKPPDRVRLPKSLFRS--LPGNRSVRLAVT 153
Db 773 IHKLPAAGATDLSFPKMGWRATGWAK---VPEDRVTVSKSVFTGTETDEAASVFVVG 829
Qy 154 ILDIGPGLFKPRILGSGVNLNRLVGLSVGMHVTKLAEPLVFSHQRPPTNMTL 213
Db 830 VLYRNLGSE----LALQNTVTLNSKVISVTV-KPPRSRLRTPLEEFAHMY---NGTT 880
Qy 214 --TCVWFVDVTKGTT-----GDSSEGCST-EVRPGETGCCDHLFFFAILLR----PT 259
Db 881 NOTCILNDETDPVSSAPPQLGPNWSRGCRTPVLDALRTRCLDRLSTFAILLASADAN 940

QY 260 LDQSTVHILTRISQACGCGVMIFLAFTIILYA-----FLRLSRERPKSEDAPKIHVALGGS 315
Db 941 MEKATLPSVTLI--VCGGVSSLTLLMLVLIYVSVWRYIRSR-----S 981
QY 316 LFLNLAF-----LVNNGSGSKGSDACWARGAVFHYFLCAFTWMCLEAFHLYLLAV 368
Db 982 VILINFCLSISSNALILGQTQTRKNVMCTLVAFLHFFFLSSFCVLTAEAWQSTMAVT 1041
QY 369 -RVNTYFGHYFLKLSLVGGLPALMV---IGTGSANSYGLYTIRDRENRSTLELCWFR- 423
Db 1042 GHLRNLIRKRELCL--GWGLPALVVAISVGFTRAKGY-----STMNYCWLSL 1087
QY 424 EHTMTALYITVHGYPFLITFLFGVMVVALVVKWIFTLSRATAVKGRKNKKVLLGLLS 483
Db 1088 EGGLEAFVGPAAAVLVNVMVIGILVFNKLVS-----DIGITDKLKRAGAS 1135
QY 484 -----SIVGVTVGLAIFTPLGLSTV---YIFALFNSLOGVFICCCWFTIL 524
Db 1136 LWSSCVVLPLLALTMWSAVLAVTDRLSALFQILFAVFDLSLEGFVIVMHCIL 1187
RESULT 2
CD97_HUMAN STANDARD; PRT; 835 AA.
ID CD97_HUMAN
AC P48960;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Leucocyte antigen CD97 precursor.
GN CD97.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95363161; PubMed=7636245;
RA Hamann J., Eichler W., Hamann D., Kerstens H.M.J., Poddighe P.J.,
Hoovers J.M.N., Hartmann J.M., Strauss M., van Lier R.A.W.;
RT "Expression cloning and chromosomal mapping of the leukocyte
RT activation antigen CD97, a new seven-span transmembrane molecule of
RT the secretion receptor superfamily with an unusual extracellular
RT domain";
RL J. Immunol. 155:1942-1950(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=ForeSkin;
RX MEDLINE=96230339; PubMed=8786105;
RA Hamann J., Hartmann E., van Lier R.A.W.;
RT "Structure of the human CD97 gene: exon shuffling has generated a new
RT type of seven-span transmembrane molecule related to the secretin
RT receptor superfamily";
RL Genomics 32:144-147(1996).
RN [3]
RP REVISIONS.
RA Hamann J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COULD BE A RECEPTOR POTENTIALLY INVOLVED IN BOTH
CC ADHESION AND SIGNALING PROCESSES EARLY AFTER LEUKOCYTE ACTIVATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD97 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd97.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

QY 516 FI 517
Db 884 FI 885

RESULT 4
BAI2_HUMAN STANDARD; PRT; 1572 AA.
ID BAI2_HUMAN
AC O60241.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 2 precursor.
GN BAI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98194217; PubMed=9533023;
RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL Cytochrome. Cell Genet. 79:103-108(1997).
CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
HEART, THYMUS, SKELETAL MUSCLE, AND DIFFERENT CELL LINES.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 GPS DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

DR EMBL; AB005298; BAA25362.1; -
DR MIM; 602683; -
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSPL.
DR InterPro; IPR001879; hormo_receptor.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp.1; 4.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00209; TSPL; 4.
DR PROSITE; PS02221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR PROSITE; PS02027; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS02026; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS00092; TSPL; 4.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Repeat.

FT SIGNAL. 1 20
FT CHAIN 21 1572
FT DOMAIN 21 924
FT TRANSMEM 925 945
FT DOMAIN 946 953
FT TRANSMEM 954 974
FT DOMAIN 975 982
FT TRANSMEM 983 1003
FT DOMAIN 1004 1024

POTENTIAL.
BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 2.
EXTRACELLULAR (POTENTIAL).
1 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
2 (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
3 (POTENTIAL).
CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1025 1045
FT DOMAIN 1046 1066
FT TRANSMEM 1067 1087
FT DOMAIN 1088 1141
FT TRANSMEM 1142 1162
FT DOMAIN 1163 1168
FT TRANSMEM 1169 1189
FT DOMAIN 1190 1572
FT TSP TYPE-1 1.
FT TSP TYPE-1 2.
FT TSP TYPE-1 3.
FT TSP TYPE-1 4.
FT GPS.
FT POLY-GLU.
FT POLY-ASN.
FT POLY-THR.
FT POLY-PRO.
FT POLY-GLY.
FT POLY-PRO.
FT CARBOHYD 94 94
FT CARBOHYD 179 179
FT CARBOHYD 180 180
FT CARBOHYD 344 344
FT CARBOHYD 425 425
FT CARBOHYD 548 548
FT CARBOHYD 633 633
FT CARBOHYD 855 855
SQ SEQUENCE 1572 AA; 171140 MW; A3775645B77BC285 CRC64;

Query Match 9.6%; Score 277; DB 1; Length 1572;
Best Local Similarity 27.6%; Pred. No. 1.1e-12;
Matches 132; Conservative 57; Mismatches 187; Indels 102; Gaps 24;

QY 118 RQVMDKDEK-PDPRVRLPKS-LFRSLPGRNSVVRLAVTI-LDIGPGTLFKGP-----RLG-169
Db 733 RRGKMDVRSDEDLPLPKVLSLSPGKATSGAAGSPGRGPGTVPVPGHSHQRL 792
QY 170 LGDGS-----GVNRLVGLSVGMHVTKLAEPLEIVFESHORPP-----PNWLT-- 214
Db 793 PADPDESSYFVIGAVLYRTGLLPLPPR-PPLAVTSRVMTVTVRPPPTQPPAEPLITVELS 851
QY 215 -----CVGMDVTK--GTTGDWSGECST-EVRPEGTCCDHLTFALLRLPTLD 261
Db 852 YIINGTDPHCASWDYSRADASGDWDTCQILETQAATRCQOHLSTFAVLAOPKD 911
QY 262 QSTVHIITRIS---QAGCGVSMIFLAFTIILY-AFLRLSRERKFSADAPKIHVALGGSIF 317
Db 912 -LTLELAGSPVPLVIGCAVSCMALLTLAIYAAFWRF-----IKSE---RSIILLNFCLS 963
QY 318 LLNLAFILVNGSGSKGSDAACWARGAVFYFLCAFTWGLEAFHLYLLAV-RVFTYFG 376
Db 964 ILASNILILVQSRVLSKGVCTMTAAFLHFFLSFSCWVLTQAMQSYLAIVGRMTRLVR 1023
QY 377 HYFLKLSLVGGLPALMV---IGTSANSYGLYTIIRDRENRSTLELCWFR-EGTTWVALY 432
Db 1024 KFLCL--GWGLPALVAVSVGFTTKGYG-----TSSVCWLSLEGGLLYARV 1069
QY 433 ITVHGFLITFLPGMVVLAIVWVKIFTLRSATAVKRGKRNKRVLTLL-----480
Db 1070 GPAAVIVLVNMLIGIIVFNKLMAR-DGISDKKKQKRGASERCPWASLLPCSAAGVPSp 1128
QY 481 GLSS-----LVGVTVGLAIETPLGLSTVY---IFALFNSLQGVFI 517
Db 1129 LLSSASARNAMASLWSSCVVPLPLALTWMSAVLWMTDRRSVLFQALFAVNSAQGEVI 1186

RESULT 5
BAI3_HUMAN STANDARD; PRT; 1522 AA.
ID BAI3_HUMAN
AC O60242; O60297;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

Query Match	8.8%	Score	254.5;	DB 1;	Length	886;
Best Local Similarity	23.2%	Pred. No.	2.7e-11;			
Matches 155; Conservative	84;	Mismatches	232;	Indels	197;	Gaps
39;						
QY	21	QBKPTGGPRNTCLGNNMYDI-----FNLDKALCFTKCROSGSDSCNVENLQRYWL---	72			
Db	274	QDPSTCGPNSICTNALGSYCCTVGFHPNPGS-----SQKDGNFSC-----QRVLKCK	323			
QY	73	-----NYEAHLMKEGLTQK-----VNTPPF--LKALVOLNST-----NTAEDEFVS	110			
Db	324	EDVIPDNKIQIOQCQEGTAVKPAYVSFCQAINNIFSVLDKVCCENKTIVSLKNTTESFPV	383			
QY	111	LEPSQVPPOVMKDED-----KPPDRVRLPKLSRSLPGNSVVRL	150			
Db	384	LQOISMTWTFTEETSSYATPLESVESMTLASFWKPSANV--TPAVRAYLDIESKVINK	442			
QY	151	AVTILDIGFTGLFGPRLG-------DGSGV-----	176			
Db	443	ECSEENVTLDLVAKDKMKIGCTTIESEESTTGTVAFVFSVGMSVLNERFFODHOAPL	502			

GCRdb; GCR_1415; .
 GCRdb; GCR_1469; .
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; hormn_receptor.
 DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 DR PROSITE; PS00227; G_PROTEIN_RECEP_F2_3; 1.
 DR PROSITE; PS00261; G_PROTEIN_RECEP_F2_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW phosphorylation.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 415 CORTICOTROPIN RELEASING FACTOR RECEPTOR
 FT 1.
 FT DOMAIN 25 121 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 122 142 1 (POTENTIAL).
 FT DOMAIN 143 151 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 152 171 2 (POTENTIAL).
 FT DOMAIN 172 189 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 190 213 3 (POTENTIAL).
 FT DOMAIN 214 227 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 228 249 4 (POTENTIAL).
 FT DOMAIN 250 268 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 269 291 5 (POTENTIAL).
 FT DOMAIN 292 314 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 315 334 6 (POTENTIAL).
 FT DOMAIN 335 349 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 350 369 7 (POTENTIAL).
 FT DOMAIN 370 415 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 415 AA; 47842 MW; 4806704B31D4C013 CRC64;
 Query Match 6.0%; Score 172.5; DB 1; Length 415;
 Best Local Similarity 24.9%; Pred. No. 1.3e-05;
 Matches 102; Conservative 57; Mismatches 178; Indels 73; Gaps 21;
 QY 143 GNRVWRL--AVTILDIGP--GTLFKGPR---LGL-GDGSVLNNRLVGLSVGQMHTVTKLA 195
 DB 2 GRPQLRLVALLLLGNPVSTSLQDORCENSLTSNVSGLOCNASVDL-IGTCWPRSPA 60
 QY 196 EPLVIFESHORPPNMTLTCTVFDVDTKGTT-----GDWSBEGCGSTVREPTGCC 245
 DB 61 GOLVV-----RCPD-----AFFGVRYNTTNGYRECLANGSWAARVNYSECQ----- 103
 QY 246 CDHITFFALLRLTLDSTVHILTRISQCGVSMIFLAFTILYFLRLSRERPKSEDA 305
 DB 104 -----EILNEEKSKVHYHVAIINYLGHICISLVALLVAFVL--FLRLSIRCLRN-- 152
 QY 306 PKHVALGSLFLNLNLAFLV---NVGSGSKGSDAACWAR--GAVPHYFLLCFAFTWGLEA 360
 DB 153 -IIHNWLSIFILRNATFWVQVLTSPVSHQSNVA-WRLVTAAYNYFHVTFWFMFEGG 210
 QY 361 PHLYLLAVRVNIFYGHYFKLSLVGWLPLAMVIGTGSANSYGLYITDRNRTSLKLC 420
 DB 211 CYLHTATVLYSTYDRLRKWN-FVCIWGVPFPIIIVAWAATKGLH-----YDN----EKC 258
 QY 421 WFRGTTMYALYITVHGYSFLITFLFGMVVALVWVKIFTLRSATVAKRGKRVKLTLL 480
 DB 259 WFGRRPGVYTDYL-YOGPMILVLLINFIENIVIRILMTKLRSITSETIQYRKAVKATL 317
 QY 481 GLSLVGVTVGLAIFPLGLSTV--YIPALFN----SLOGVFCICWFTIL 524
 DB 318 VLUPLLGITITMLFVNP-GEDEVSRVVFTYFNFSFLESFQGFVSVFYCF 366
 RESULT 12
 DIHR_ACHDO STANDARD; PRT; 441 AA.
 ID DIHR_ACHDO Q16983;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Diuretic hormone receptor precursor (DH-R).
 OS Acheta domestica (House cricket).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Insecta; Orthoptera; Orthopteroidea; Gryllidae;
 OC Gryllinae; Acheta.
 OX NCBI_TaxID:6997;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Malpighian tubules;
 RX MEDLINE=96189577; PubMed=8673074;
 RA Reagan J.D.;
 RT "Molecular cloning and function expression of a diuretic hormone
 receptor from the house cricket, Acheta domestica.";
 RL Insect Biochem. Mol. Biol. 26:1-6(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE INSECT DIURECTIC HORMONE. THE ACTIVITY
 OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL
 CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MALPIGHIAN TUBULES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation ~
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/annoucnce/
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U15959; AAC47000.1; .
 DR GCRdb; GCR_1380;
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR001879; hormn_receptor.
 DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF02793; HRM; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00008; Hormr; 1.
 DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
 DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
 DR PROSITE; PS00227; G_PROTEIN_RECEP_F2_3; 1.
 DR PROSITE; PS00261; G_PROTEIN_RECEP_F2_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 441 DIURECTIC HORMONE RECEPTOR.
 FT DOMAIN 25 134 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 135 158 1 (POTENTIAL).
 FT DOMAIN 159 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 187 2 (POTENTIAL).
 FT DOMAIN 188 194 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 195 224 3 (POTENTIAL).
 FT DOMAIN 225 238 4 (POTENTIAL).
 FT TRANSMEM 239 260 5 (POTENTIAL).
 FT DOMAIN 261 291 6 (POTENTIAL).
 FT TRANSMEM 292 315 7 (POTENTIAL).
 FT DOMAIN 316 338 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 339 357 8 (POTENTIAL).
 FT DOMAIN 358 371 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 372 391 9 (POTENTIAL).
 FT DOMAIN 392 441 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 441 AA; 49538 MW; 501915AC2E776CSC CRC64;

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 22:24:36 ; Search time 68.53 Seconds
(without alignments)
1385.878 Million cell updates/sec

Title: US-09-733-387-44

Perfect score: 2878

Sequence: 1 MATPRGLGALLLLLLLPTSG.....STTVSSSTARLDQAHSASQE 549

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rhodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_ivirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	643.5	22.4	687	11	Q9QZT2
2	610.5	21.2	1014	4	Q00406
3	603.5	21.0	687	4	Q05966
4	596.5	20.7	693	4	Q96HB4
5	594.5	20.7	693	4	Q9Y653
6	555.5	19.3	512	4	Q96JW0
7	555.5	19.3	541	4	Q9Y3K0
8	387	13.4	872	4	Q94867
9	383.5	13.3	1231	6	Q97819
10	383.5	13.3	1274	6	Q97820
11	383.5	13.3	1299	6	Q97825
12	383.5	13.3	1342	6	Q97826
13	383.5	13.3	1503	6	Q97818
14	383.5	13.3	1527	11	O88927
15	383.5	13.3	1571	6	Q97824
16	382.5	13.3	823	4	Q9UHX3

17	382.5	13.3	1240	4	Q9HAR2
18	379	13.2	1240	6	Q97822
19	379	13.2	1283	6	Q97823
20	379	13.2	1308	6	Q97828
21	379	13.2	1351	6	Q97829
22	379	13.2	1512	6	Q97821
23	379	13.2	1550	11	Q9Z173
24	379	13.2	1580	6	Q97827
25	371.5	12.9	839	4	Q96IE7
26	371.5	12.9	1469	4	Q9HAR3
27	371.5	12.9	1474	4	Q94910
28	370.5	12.9	1467	6	Q97830
29	370.5	12.9	1472	6	Q97831
30	365.5	12.7	1515	11	O88917
31	361.5	12.6	1021	4	Q94882
32	359.5	12.5	1354	6	Q97806
33	359.5	12.5	1397	6	Q97807
34	359.5	12.5	1420	6	Q97814
35	359.5	12.5	1463	6	Q97815
36	355.5	12.4	1478	11	O88923
37	355	12.3	1123	4	Q9UKY5
38	355	12.3	1177	4	Q9UKY6
39	355	12.3	1403	4	Q95490
40	353	12.3	1341	6	Q97802
41	353	12.3	1384	6	Q97803
42	353	12.3	1407	6	Q97810
43	353	12.3	1450	6	Q97811
44	351	12.2	606	4	Q9HBW9
45	351	12.2	1487	11	Q9Z174

ALIGNMENTS

RESULT 1

Q9QZT2	ID	Q9QZT2	PRELIMINARY;	PRT;	687 AA.
AC	Q9QZT2:				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	SERPENTINE RECEPTOR.				
GN	GPR56 OR CYT28.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SYRAIN-C57BL/6; TISSUE=LIVER;				
RA	Phillips R.L., Ernst R.E., Bosil M., Wesley C.K., Moore K.A.,				
RA	Kingsley P.D., Sykes S., Palis J., Lemischka I.R.;				
RT	"Identification of novel hematopoietic stem cell regulatory genes."				
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF166382; AAF00617.1;				
DR	MGD; MGI:1340051; Gpr56				
DR	InterPro; IPR000873; AMP-bind.				
DR	InterPro; IPR000832; GPCR_secretin.				
DR	InterPro; IPR000203; PKD_cys_rich.				
DR	Pfam; PF00002; 7tm_2; 1.				
DR	Pfam; PF01825; GPS; 1.				
DR	PRINTS; PR00249; GPCRSECRETIN.				
DR	SMART; SM00303; GPS; 1.				
DR	PROSITE; PS00455; AMP_BINDING; UNKNOWN_1.				
DR	PROSITE; PS0261; G_PROTEIN_RECP_F2_4; 1.				
KW	Receptor.				
SQ	SEQUENCE	687 AA;	77271 MW;	B5315D70AF443809	CRC64;

Query Match 22.4%; Score 643.5; DB 11; Length 687;
Best Local Similarity 31.1%; Pred. No. 1.2e-47;
Matches 168; Conservative 89; Mismatches 219; Indels 65; Gaps 17;


```

QY 269 TRISOAGCGVMIFLAFTIILYAFRLSRERFKSEADAPKIRHVALGGSLFLINLAFLVNG 328
Db 626 TFITYIGCGSLSFSLVLTIVY----IAFKIRRDYPSKILLQICALLLNLFVLDLSDW 681
QY 329 SSGKSDAACWARGAVHYFLLCATWGLAEAFHYLLAVRVNTYFGHYFLKLSLVGNG 388
Db 682 IALYKMOGICISVAVFLHYFLVSTWGLAEAFHYLLAVRVNTYIRKYILKFCIVGNG 741
QY 389 LPALMV--IGTGSANSYGLYTRDRENRSTSLCWCWREGTMTYALXYTVHGYPFLITFLG 446
Db 742 VPVVVVTIILTSPDNYGLSGVKGKPNPDDFCWINNNA---VFYITVGVFCVFIEN 798
QY 447 MVVALVWVKITLSRATAVAKRGNKKVL-----TLGLSSLVGVWGLAIFT--PIGL 500
Db 799 VSMFIWV--LVQLCRIRKKKLGQAQRKTSIQDLRSIAGLFLGLITWGAFFAWGPNV 855
QY 501 STVYIFALNSLQGVFCOWETI 523
Db 856 TFWLFAIENTLQGFIFIFYCV 878

RESULT 3
O95966 PRELIMINARY; PRT; 687 AA.
ID AC O95966
AC O95966;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TM7XN1 PROTEIN PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99198960; PubMed=10100861;
RA Zendenman A.J.W., Cornelissen I.M.H.A., Weidie U.H., Ruiter D.J.,
RA Van Muijen G.N.P.;
RT "TM7XN1, a novel human EGF-TM7 like protein, detected with mRNA
RT differential display using human melanoma cell lines with different
RT metastatic potential";
RL FEBS Lett. 446:292-298(1999).
DR EMBL; AJ011001; CAB37294.1; -.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR000832; GPCR_secretin.
DR Pfam; PF00002; 7tm2; 1.
DR Pfam; PF01825; GFS; 1.
DR PRINTS; PRO0249; GPCRSECRETIN.
DR SMART; SM00303; GFS; 1.
DR PROSITE; PS00455; AMP_BINDING; UNKNOWN_1.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
KW SIGNAL.
FT SIGNAL.
FT CHAIN.
FT CHAIN.
SQ SEQUENCE 687 AA; 77150 MW; B3F0EE75B0CD5CF1 CRC64;

Query Match 21.0%; Score 603.5; DB 4; Length 687;
Best Local Similarity 34.9%; Pred. NO. 3.6e-44;
Matches 149; Conservative 66; Mismatches 177; Indels 35; Gaps 10;

QY 131 VRLPKSLFRSLPGRNSRVRLAVTILDIGPTLFKGPRLGDSGVLNRLVGLSVGMH 190
Db 267 VLLPRTLFTQTKRGRGEAEKRLLLVDFSSQALFQDK-----NSSHVIGKEKVLGIVVQNTK 321
QY 191 VTKLAEPLFIVFQHPPPNMTITCVFW--DVTGTTGDSWSEGGSTEVREPGTVCCDDH 248
Db 322 VANLTPVVLTFQHQLOPKNVTLCQVFWVEDPTLSSPGHWSAGCETVRRRETQTSFCNH 381
QY 249 LTFALLRLRLDQSTV--HILTRISQACGVMIFLAFTIILYAFRL--LSRERFKSED 304
Db 382 LTYFAVLWSSVEVDVAVHKHYSLLSYGCVVVSALACLVTTAAYLCSRVLPCCRKKPRDY 441
QY 305 APIKHALGSLFLNLAFVWNGSGSKSDAACWARGAVHYFLLCATWGLAEAFHY 364
Db 442 TIKVHNLLAVFLDTSFLSEPVALGSEAGCRASALFLHLSLTCLSMWGLEGINLY 501
QY 365 LLAVRVNTYFGHYFLKLSLVGNGPALMV--IGTGSANSY--LYTIRDRNRTSL 419
Db 502 RLVEVEFTVYVPGYLLKLSAMGWPFIETLVVALVDVNDYGPILAVHRTPEGVYPSM 561
QY 420 CWFREGTMTYALXYTVHGYPFLITFLGVMVVALVWVKITLSRATAVAKRGNKKVL 479
Db 562 CWRDLSLV--YITWGLFSLVFLFNMAVLATMVQILRLPHT-----QKSHVLT 612
QY 480 LGLSSLVGVWGLAIFT----PLGLSTVYIFALNSLQGVFCOWETI 531

```

```

QY 307 KIHVALGSLFLNLAFVWNGSGSKSDAACWARGAVHYFLLCATWGLAEAFHYLL 366
Db 438 KYHNNLLAVFLDTSFLSEPVALGSEAGCRASALFLHLSLTCLSMWGLEGINLYRL 497
QY 367 AVRVNTYFGHYFLKLSLVGNGPALMV--IGTGSANSY--LYTIRDRNRTSL 421
Db 498 VVEVEFTVYVPGYLLKLSAMGWPFIETLVVALVDVNDYGPILAVHRTPEGVYPSM 557
QY 422 FREGTMTYALXYTVHGYPFLITFLGVMVVALVWVKITLSRATAVAKRGNKKVL 481
Db 558 IRDLSLV--YITWGLFSLVFLFNMAVLATMVQILRLPHT-----QKSHVLT 608
QY 482 LSLVGVWGLAIFT----PLGLSTVYIFALNSLQGVFCOWETI 533
Db 609 LSLVGLPWALLFFSFASGTFOVLVYLFIITTSFQGLFIETWYMSRLQARGSPSLK 668
QY 534 SSSARL 540
Db 669 NSDSARL 675

RESULT 4
O96HB4 PRELIMINARY; PRT; 693 AA.
ID AC O96HB4
AC O96HB4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE G PROTEIN-COUPLED RECEPTOR 56.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-PLACENTA, AND CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008770; AA08770.1; -.
KW Receptor.
SQ SEQUENCE 693 AA; 77746 MW; 5C9FA8194D1B5C50 CRC64;

Query Match 20.7%; Score 596.5; DB 4; Length 693;
Best Local Similarity 34.7%; Pred. NO. 1.5e-43;
Matches 149; Conservative 66; Mismatches 181; Indels 33; Gaps 10;

QY 131 VRLPKSLFRSLPGRNSRVRLAVTILDIGPTLFKGPRLGDSGVLNRLVGLSVGMH 190
Db 267 VLLPRTLFTQTKRGRGEAEKRLLLVDFSSQALFQDK-----NSSHVIGKEKVLGIVVQNTK 321
QY 191 VTKLAEPLFIVFQHPPPNMTITCVFW--DVTGTTGDSWSEGGSTEVREPGTVCCDDH 248
Db 322 VANLTPVVLTFQHQLOPKNVTLCQVFWVEDPTLSSPGHWSAGCETVRRRETQTSFCNH 381
QY 249 LTFALLRLRLDQSTV--HILTRISQACGVMIFLAFTIILYAFRL--LSRERFKSED 304
Db 382 LTYFAVLWSSVEVDVAVHKHYSLLSYGCVVVSALACLVTTAAYLCSRVLPCCRKKPRDY 441
QY 305 APIKHALGSLFLNLAFVWNGSGSKSDAACWARGAVHYFLLCATWGLAEAFHY 364
Db 442 TIKVHNLLAVFLDTSFLSEPVALGSEAGCRASALFLHLSLTCLSMWGLEGINLY 501
QY 365 LLAVRVNTYFGHYFLKLSLVGNGPALMV--IGTGSANSY--LYTIRDRNRTSL 419
Db 502 RLVEVEFTVYVPGYLLKLSAMGWPFIETLVVALVDVNDYGPILAVHRTPEGVYPSM 561
QY 420 CWFREGTMTYALXYTVHGYPFLITFLGVMVVALVWVKITLSRATAVAKRGNKKVL 479
Db 562 CWRDLSLV--YITWGLFSLVFLFNMAVLATMVQILRLPHT-----QKSHVLT 612
QY 480 LGLSSLVGVWGLAIFT----PLGLSTVYIFALNSLQGVFCOWETI 531

```

Db 613 LGLSLVLGFWALIFFSFASGTQLVLYLFSITTSFQGLFIWIWWSMRLOARQGPSPL 672
QY 532 TVSSSTARL 540
Db 673 KNSDSARL 681

RESULT 5
ID Q96JW0 PRELIMINARY; PRT; 693 AA.
AC Q96JW0;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE G-PROTEIN-COUPLED RECEPTOR.
GN GPR56.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=99168899; PubMed=10049584;
RX Liu M., Parker R.M., Darby K., Eyre H.J., Copeland N.G., Crawford J.,
RA Gilbert D.J., Sutherland G.R., Jenkins N.A., Herzog H.;
RT "GPR56, a novel secretin-like human G-protein-coupled receptor gene.";
RL Genomics 55:296-305(1999);
DR EMBL; AF106858; AAD30545.1; .
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR SMART; SM00303; GPS; 1.
DR PROSITE; PS00455; AMP_BINDING; UNKNOWN_1.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
KW Receptor.
SQ SEQUENCE 693 AA; 77753 MW; F71C8E62666A5141 CRC64;

Query Match 20.7%; Score 594.5; DB 4; Length 693;
Best Local Similarity 34.7%; Pred. No. 2.3e-43;
Matches 149; Conservative 65; Mismatches 182; Indels 33; Gaps 10;
QY 131 VPLPKSLFSLRSGNSVRLAVTILDIGPGLFGPRGLGDSGLNRLNRLVGLSGQM 190
Db 267 VLLPRTLFTQKRGSGEAEKRLLLVDFSSQALFQK-----NSSQVLGEKVLGVVQNTK 321
QY 191 VTKLAEPLEIVFHSQRPNNMTLTCVFW--DVTGTTGDMSEGCSTEVREPGTVCCDDH 248
Db 322 VANLPEPVLTFOHLQPKNTVLCQVFWEDPTLSSPGHWSAGCEVRRRETQTSFCFNH 381
QY 249 LTFALLRLPTDQSTV--HILTRISQAGCGVMIFLAFTILYLAFLR--LSRERFKSED 304
Db 382 LTYFAVLVSSVEVDVAKHKLSSLSSYGCYVSALACLVITAAYLCSRVLPFCRRKPRDY 441
QY 305 APKIHVALGGSFLNLNLAFLNVGSGSDAACWARGAFHYFLLCFTWGLEAFHLY 364
Db 442 TKVHNNLLAVFLDTSFLSEPALTGSEAGCRASAFILHLSLTLCLSWMGLEGYNLY 501
QY 365 LLAARVNTYFGHYFLKLSLVGWSGLPALMV--IGTGSANSYG---LYTIRDENRTSLEL 419
Db 502 RLWVEVFQTPGYLLKLSAMGWGPFIFLTVLVALVDVNDYGPITLAVHRTPEGVYPSM 561
QY 420 CNFRETTMYALYIVHGYFLTLTFGMVLALVWVKITLSRATAVAKERKRRKVLTL 479
Db 562 CWIRSLVS---YTNLGLFSLVFLFNMAMLATWVQILRLRPHI-----QKWSHVLTL 612
QY 480 LGLSSLVGVTWGLAIFT-----PLGSTVYVIFALNSLOGVFCCTFTLYLPDSO----T 531
Db 613 LGLSLVLGFWALIFFSFASGTQLVLYLFSITTSFQGLFIWIWWSMRLOARQGPSPL 672

QY 532 TVSSSTARL 540
Db 673 KNSDCARL 681
RESULT 6
ID Q96JW0 PRELIMINARY; PRT; 512 AA.
AC Q96JW0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDNA FLJ14937 FIS, CLONE PLACE1010231L, WEAKLY SIMILAR TO CELL
DE SURFACE GLYCOPROTEIN EMR1 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA ISOGAI T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Watanabe S., Hosoi T., Raku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Ninomiya K., Iwavanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027843; BAB55406.1; .
SQ SEQUENCE 512 AA; 57870 MW; BB4EC50BF3FCDDBE CRC64;

Query Match 19.3%; Score 555.5; DB 4; Length 512;
Best Local Similarity 33.7%; Pred. No. 3.9e-40;
Matches 140; Conservative 66; Mismatches 161; Indels 49; Gaps 13;
QY 127 PPDVRVLPKSLFSLRSGNSVRLAVTILDIGPGLFGPRGLGDSGLNRLNRLVGLSV 186
Db 10 PLASVILPPNLENLSPEDSVLVRRAQFTFFNKTLGQ----DVGPRKTLVSYMACSI 65
QY 187 GQHVTKLAEPLEIVFHSQRPNNMTLTCVFWNDVTGTT--GDWSSSGCSTEVREPG--TV 243
Db 66 GNITQNLKDPVOIKTKHTRTQEVHHPICAFWDLNKNKSGGNTSGCVAHDRSDASEIV 125
QY 244 CCDDHLTFFALL-----RPTLDQSTVHILTRISQAGCGVMIFLAFTILY--AFURLSR 297
Db 126 CLCNHFTFGVMDLPRASQDLARNTKVLTFISYICGIGISAIFSAATLLTVVAFELRR 185
QY 298 ERPKSEDAP--KIHVALGGSFLNLNLAFLNVGSGSDAACWARGAFHYFLLCFTW 356
Db 186 -----DYPKILMNLSTALLFLNLFLDGTITPSFNDGLCIAVAVLHFFLLAFTTWM 239
QY 357 GLFAFLYLLAARVNTYFGHYFLKLSLVGWSGLPALMV--IGTGSANSYGILYITIRDENRT 415
Db 240 GLCAHMTIALVKNVNTYIRYILKFCIIGWGLPALVSVVSLASRNNNEVYKESYGEK 299
QY 416 SLELCNFRGTTMYALYIVHGYFLTLTFGMVLALVWVKITLSRATAVAKERKN---472
Db 300 GDEFQWIDPV---IFYVTCAGYFGVMFFLNIAFMIVMVQIC-----GRNGR 345
QY 473 -----RKVL-----TLGLSLVGVTVWGLAIFT--PLGSTVYVIFALNSLOGVFI 517
Db 346 SNRTLREEVLRNLRKSVSVLFLGWTGFAFFAWGPNLNPFFMYLFSIFNSLQGLFI 401
RESULT 7
ID Q9Y3K0 PRELIMINARY; PRT; 541 AA.
AC Q9Y3K0;
DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE DJ287G14.2 (PUTATIVE NOVEL SEVEN TRANSMEMBRANE DOMAIN PROTEIN)
 DE (FRAGMENT).
 GN DJ287G14.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mashreghi-Mohammadi M.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL033377; CAB42901.1; -.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR000203; PKD_cys_rich.
 DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF01825; GPS; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00303; GPS; 1.
 DR PROSITE; PS00249; GPCRSECRETIN.
 DR PROSITE; PS0261; G_PROTEIN_RECEPTOR_F2_4; 1.
 KW Transmembrane.
 FT NON_TER
 SQ SEQUENCE 541 AA; 61048 MW; 0672C54BEEC7E91D CRC64;
 Query Match 19.3%; Score 555.5; DB 4; Length 541;
 Best Local Similarity 33.7%; Pred. No. 4.2e-40;
 Matches 140; Conservative 66; Mismatches 161; Indels 49; Gaps 13;
 QY 127 PDRVRLPKSLRSPGNRSVRLAVTILDIGPGLFGRPLGLDGSGLNNRLVGLSV 186
 Db 10 PLASVILPNNLENSPDSVLVRRQAQTFNKTGLFQ----DVGPKRTLVSYVMACSI 65
 QY 187 GQHVTKLAEPLEIVESHORPPNMLTLCVFDVTKGTT-GDWSSEGCSTVRPEG--TV 243
 Db 66 GNITQNLDPQVQIKTKHTQEVHPHICAFWDLNKNKSFQGNWTSQCAVHRDSDASETV 125
 QY 244 CCDDHLTFALLL-----RPTLDQSTVHILTRISQAGGVSMIFLAFTIILY-AFLRLSR 297
 Db 126 CLCNHFTFGVLMPLPSASQLDARTKVLTFISYIGCGISAIFSAATLLTYVAEKLRR 185
 QY 298 ERKSEDAK-KHVALGGSFLFLNLAFLVNVGSGSGSDACWARGAVHFELCAFTWM 356
 Db 186 -----DYPKSKILMNLSTALLFLNLLFDLGDWITSFNDVGLCIAVAVLLHFLFLATFTWM 239
 QY 357 GLEAFHLYLAVRVNTYFGHYFLKLSLVGGLPALMV-IGTGSANSYGLYTIIRDRENT 415
 Db 240 GLEAHMYIALVKVNTYIRRYILKECIIGWGLPALVSVVSLASRNNNEVYKESYGEK 299
 QY 416 SLELCWFREGTMYALYIVHGYFLITFLFGMVVLAFLVVKVIFTLRSATAVKERGKN--- 472
 Db 300 GDFECWTDQDV---IFYVTGAGYFGYMFNLAMFIVVMVQIC-----GRNGKR 345
 QY 473 -----RKVL-----TLGLSSLVGVTVGLAIFT--PLGLSTVYIFALFNSLQGVFI 517
 Db 346 SNRTLEEVLRNLSVVSLTFELGLTGWGPAFWAGPLNIPFMYLSIENSLOGLFI 401
 RESULT 8
 ID O94867 PRELIMINARY; PRT; 872 AA.
 AC O94867;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE KIAA0768 PROTEIN (FRAGMENT).
 GN KIAA0768.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=99087487; PubMed=9872452;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 5:277-286(1998).
 DR EMBL; AB018311; BAA34488.1; -.
 DR InterPro; IPR000832; GPCR_secretin.
 DR InterPro; IPR003334; Latrophilin.
 DR InterPro; IPR000203; PKD_cys_rich.
 DR Pfam; PF00002; 7tm_2; 1.
 DR Pfam; PF01825; GPS; 1.
 DR Pfam; PF02354; Latrophilin; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00303; GPS; 1.
 DR PROSITE; PS00249; GPCRSECRETIN.
 DR PROSITE; PS0261; G_PROTEIN_RECEPTOR_F2_4; 1.
 KW NON_TER
 FT NON_TER
 SQ SEQUENCE 872 AA; 97420 MW; 9689996CEBF3BD46 CRC64;
 Query Match 13.4%; Score 387; DB 4; Length 872;
 Best Local Similarity 27.1%; Pred. No. 4.2e-25;
 Matches 137; Conservative 84; Mismatches 198; Indels 86; Gaps 25;
 QY 59 SDSNVENLQRYWLYEAHLMKEGLTQ----KVNTPFLLKALVQNLSTNATDFFSLEPS 114
 Db 72 SDQLRAATMLLTVEESAFVADNLLKTDIVRENTDNKLEVARLSTEG-----NLEDL 125
 QY 115 QVPQVMKDEKPPDRVRLPKSLRSPGNRSVRLAVTIL-DIGPGLFGRPLGLG-- 171
 Db 126 KPEPNMGH-----STIQLSANTLQK-NGRNGEIRVAFVLYNNLGPYLSTENASKMLGTE 179
 QY 172 ----DGSGLVNLRLVGLSVGQMHVTK--LAEPLEIVFVSH-ORPPNMLTLCVFDVTKGT 224
 Db 180 ALSTNHSVIVNSPVITAAINKFESNKKVYADPVVTFVTHIKQSEENFNPNCSFWSYKRT 239
 QY 225 -TGDMSEGCs-TEVRPEGTGCCDHLAFFALLR--PTLDQSTVH--ILTRISOAGCV 278
 Db 240 MTGYNSTQGRLLTNTKHTTCSNHLTNFVLMHVEVVKHSDAVHDLDDVITWVGILL 299
 QY 279 SMFLAFTIILAFRLSRERKSEDAKPHIHALGGSFLFLNLAFLVNVGSGSGSDAAC 338
 Db 300 SLVCLLICITFCFF-----RGLQSDRNTIHKNLICISLFAELJEL--IGINRTDQPTAC 352
 QY 339 WARGAVHFYLLCAFTWGLFAFLYLLAVRVNTYFGH--YFLKLSLVGGLPALMVIG 396
 Db 353 AVFAALLHFEFFLAFTWMELEGVQLYIMLVEFSEHSRRKYFY--LVGYGMPALIVAV 409
 QY 397 TGSAN--SYGLYTIIRDRENTSLELCWFREGTMYALYIVHGYFLITF-----LFG 446
 Db 410 SAADVRSYG-----TDKVCWLRDT-----YFINSFIGPATLIIMLN 447
 QY 447 MYVALVWVKIFTLSRATAV--KEGG--KNRKK-VTLGLSSLVGVTV--GLAIFTPLG 499
 Db 448 VIFGLTALYMP---HWTAILPESCCLDNIKSWIGALICLLGLTWAGLWYINEST 504
 QY 500 LSTVYIFALFNSLQGVFICCNFTIL 524
 Db 505 VMAYLFTIFNSLQGMFIFHCVL 529
 RESULT 9
 ID O97819 PRELIMINARY; PRT; 1231 AA.
 AC O97819;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE LATROPHILIN 3 SPLICE VARIANT ABAG.
 OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148828; PubMed=10025961;
RA Matsushita H., Lelianova V.G., Ushkaryov Y.A.;
RT "The latrophilin family: multiply spliced G protein-coupled receptors
with differential tissue distribution.";
RL FEBS Lett. 443:348-352(1999).
DR EMBL: AF11086; AAD05322.1; .
DR InterPro; IPR000922; Gal_lectin.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003334; Latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF02140; Gal_lectin; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR ProDom; PD005612; Gal_lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS0227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS0228; SUEL_LECTIN; 1.
SQ SEQUENCE 1231 AA; 138200 MW; 1258A8A05D64325D CRC64;

Query Match 13.3%; Score 383.5; DB 6; Length 1231;
Best Local Similarity 27.6%; Pred. No. 1.3e-24;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;
QY 77 HLMKEGLTKVNTFFLKALVQLST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
DB 683 NLLKTDIVRE-NTDNQLEVARLSTEGNLEDKFPENTGHSTIQLSANTLKN----- 735
QY 131 VRLPKSLFRSLPGNRSVRLAVTIL-DIGPGTLFGPRLGLG-----DGSGLVNNRLVG 183
DB 736 -----GRNGEIRVAVLYNGLPYLSTENASMKLGTTEAMSTNHSVIVNSPVIT 783
QY 184 LSVGMQHVTK--LAEPLEIVESH-QRPPNMTLCVFDVTKGT-TGDSSEGCS-TEVR 238
DB 784 AAINKEFSNKVYLDPPVFTVKHKQSEENFNPCSWFSYKRTMTGYWSTQGCRLTTN 843
QY 239 PEGTVCCDHLTFFALLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
DB 844 KTHHTCSNHLTNFAVMAHVEVKHSDAHDLLDVITWVIGLLSLVCLLICITFCFF- 902
QY 295 LSREFKSEDPKIHVALGSLFLNLAFVNVGSGKSDAACWARGAVHYFLCAFT 354
DB 903 ----RGLQSDRNTIHKNCISLFAELLFL--IGINTDQPIACAVFAALLHFFFLAAFT 956
QY 355 WMGLEAFHLYLLAVRVNTYFGH--YFLKLSLVGNGLPALMVGITGSAN--SYGLYTIRD 410
DB 957 WMFLEGVQLYIMLVFEFSEHSRRKYFY--LVVGYMPALIVASAAVDYRSYG----- 1007
QY 411 RENRTSLELCWFREGTMYALYIVVHGYFLITF-----LFGMVVALVVMVIFLISR 462
DB 1008 -----TDKVCWLRLDT-----YFIWSFIGPATLIIMNLNIFLGIALYKMF---H 1048
QY 463 ATAV--KERG--KNRKK-VLTLLGLSSLVGVTW--GLAIFTPGLSTVYIFALNSLQGV 515
DB 1049 HTAILKPESGCDNCKSWVIGAILLCLLGLTWAFGLMYINESTVIMAYLFTIFNSLQGM 1108
QY 516 FICCCWFTIL 524
||| :| :

DB 1109 FIFIEHCVL 1117
RESULT 10
ID O97820 PRELIMINARY; PRT; 1274 AA.
AC O97820;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE LATROPHILIN 3 SPLICED VARIANT ABAB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148828; PubMed=10025961;
RA Matsushita H., Lelianova V.G., Ushkaryov Y.A.;
RT "The latrophilin family: multiply spliced G protein-coupled receptors
with differential tissue distribution.";
RL FEBS Lett. 443:348-352(1999).
DR EMBL: AF11087; AAD05323.1; .
DR InterPro; IPR000922; Gal_lectin.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003334; Latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF02140; Gal_lectin; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02354; Latrophilin; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR ProDom; PD005612; Gal_lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS0227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS0228; SUEL_LECTIN; 1.
SQ SEQUENCE 1274 AA; 142748 MW; 147C1B5FC160037F CRC64;
Query Match 13.3%; Score 383.5; DB 6; Length 1274;
Best Local Similarity 27.6%; Pred. No. 1.4e-24;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;
QY 77 HLMKEGLTKVNTFFLKALVQLST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
DB 683 NLLKTDIVRE-NTDNQLEVARLSTEGNLEDKFPENTGHSTIQLSANTLKN----- 735
QY 131 VRLPKSLFRSLPGNRSVRLAVTIL-DIGPGTLFGPRLGLG-----DGSGLVNNRLVG 183
DB 736 -----GRNGEIRVAVLYNGLPYLSTENASMKLGTTEAMSTNHSVIVNSPVIT 783
QY 184 LSVGMQHVTK--LAEPLEIVESH-QRPPNMTLCVFDVTKGT-TGDSSEGCS-TEVR 238
DB 784 AAINKEFSNKVYLDPPVFTVKHKQSEENFNPCSWFSYKRTMTGYWSTQGCRLTTN 843
QY 239 PEGTVCCDHLTFFALLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
DB 844 KTHHTCSNHLTNFAVMAHVEVKHSDAHDLLDVITWVIGLLSLVCLLICITFCFF- 902
QY 295 LSREFKSEDPKIHVALGSLFLNLAFVNVGSGKSDAACWARGAVHYFLCAFT 354
DB 903 ----RGLQSDRNTIHKNCISLFAELLFL--IGINTDQPIACAVFAALLHFFFLAAFT 956
QY 355 WMGLEAFHLYLLAVRVNTYFGH--YFLKLSLVGNGLPALMVGITGSAN--SYGLYTIRD 410
||| :| :

```
Db 957 WMFEGVQLYIMLVFESEHSRRKYFY---LVGYGMPALIVAVSAADYRSYG----- 1007
QY 411 RENRTSLCWFREGTMYALYITVHGXYLTF-----LFGMVVLALVVKWKTLSR 462
Db 1008 -----TDKVCWLRDLT-----YFWSFGPATLIIMLVIFGLIYKMF---H 1048
QY 463 ATAV--KERG--KNRKK-VLTLLGSSLVGVTV--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1049 HTAILKPESGCLDNKISWIGAIALLCGLTWAFGLMYNESTVIMAYLFTIFNSLQGM 1108
QY 516 FICCWFTIL 524
Db 1109 FIFIFHCVL 1117

RESULT 11
097825 PRELIMINARY; PRT; 1299 AA.
AC 097825;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE LATROPHILIN 3 SPLICED VARIANT BBAG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148828; PubMed=10025961;
RA Matsushita H., Lelianova V.G., Ushkaryov Y.A.;
RT "The latrophilin family: multiply spliced G protein-coupled receptors
with differential tissue distribution.";
RL FEBS Lett. 443:348-352(1999).
DR EMBL; AF111092; AAD05328.1; -
DR InterPro; IPR000922; Gal_lectin.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003334; Latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm2; 1.
DR Pfam; PF02140; Gal_lectin; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02354; Latrophilin; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR ProDom; PD005612; Gal_lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS50228; SUEL_LLECTIN; 1.
SQ SEQUENCE 1299 AA; 145109 MW; 9D83D9258D7FA821 CRC64;

Query Match 13.3%; Score 383.5; DB 6; Length 1299;
Best Local Similarity 27.6%; Pred. No. 1.4e-24;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

QY 77 HLMKEGLTQKVTNPFKALVQNLST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
Db 751 NLKLTDIVRE-NTNDNIQLEVARLSTEGNLEDLKFPENTGHSTIQLSANTLKQN----- 803
QY 131 VRLPKSLFRSLPCNRSVRVLAVTIL-DIGPGTLFKGPRGLG-----DGSGVLNRLVG 183
Db 804 -----GRNGEIRVAFVLYNNLGPYLSTENASMKLGTTEAMSTNHSVIVNSPVIT 851
QY 184 LSYGQMHVTK--LAEPLEIVFSH-QRPPPNMTLTCVFDVTKGT-TGDWSSEGCs-TEVR 238
::: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|:: |::|::
```

```
Db 852 AAINKEFSNKVYLADPVVFTVKHIKQSEENFNPCNCSFWSYSKRTMTGYWSTQGCRLTTN 911
QY 239 PEGVVCCHDLTFEALLR--PTLDQSTVH--ILTRISOAGCGVSMIFLAFTIILYAFLR 294
Db 912 KHTTCSGNHLTNFAVLMAHVEVKHSDAHDLLDVLTVWGIILSLVCLICITFCFF- 970
QY 295 LSRRFKSEDAKTHVALGGSLFLNLAFLVNMGSSGSGSDAACWARGAVHYFLCAFT 354
Db 971 ---RGLQSDRNTIHKNLCTSLFVAELLFL--IGINRTDQPIACAVEAALLHFFFLAAFT 1024
QY 355 WMGLEAFHLYLLAVRVNTYFGH--YFLKLSLVGWGLPALMVIGTGSAN--SYGLYIRD 410
Db 1025 WMFEGVQLYIMLVFESEHSRRKYFY---LVGYGMPALIVAVSAADYRSYG----- 1075
QY 411 RENRTSLCWFREGTMYALYITVHGXYLTF-----LFGMVVLALVVKWKTLSR 462
Db 1076 -----TDKVCWLRDLT-----YFWSFGPATLIIMLVIFGLIYKMF---H 1116
QY 463 ATAV--KERG--KNRKK-VLTLLGSSLVGVTV--GLAIFTPLGLSTVYIFALFNSLQGV 515
Db 1117 HTAILKPESGCLDNKISWIGAIALLCGLTWAFGLMYNESTVIMAYLFTIFNSLQGM 1176
QY 516 FICCWFTIL 524
Db 1177 FIFIFHCVL 1185

RESULT 12
097826 PRELIMINARY; PRT; 1342 AA.
AC 097826;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LATROPHILIN 3 SPLICED VARIANT BBAG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148828; PubMed=10025961;
RA Matsushita H., Lelianova V.G., Ushkaryov Y.A.;
RT "The latrophilin family: multiply spliced G protein-coupled receptors
with differential tissue distribution.";
RL FEBS Lett. 443:348-352(1999).
DR EMBL; AF111093; AAD05329.1; -
DR InterPro; IPR000922; Gal_lectin.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003334; Latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm2; 1.
DR Pfam; PF02140; Gal_lectin; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02354; Latrophilin; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR ProDom; PD005612; Gal_lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS50228; SUEL_LLECTIN; 1.
SQ SEQUENCE 1342 AA; 149658 MW; 51FBC79F37FF5B4B CRC64;

Query Match 13.3%; Score 383.5; DB 6; Length 1342;
Best Local Similarity 27.6%; Pred. No. 1.5e-24;
```

```
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

QY 77 HLMKEGLTKVNTPELKALVQLNST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
D 751 NLLKTDIVRE-NTNIOLEVARLSTEGNEDLKFPENTGHSTIQLSANTLKQN-----803
QY 131 VRLPKSLFRSLPGNRSVRLAVTIL-DIGPGTLFKGPRGLG-----DGSGLVNLRLVG 183
D 804 -----GRNGEIRVAFVLYNNGPYLSTENASMKLGTEAMSTNHSVIVNSPVIT 851
QY 184 LSVGQMHVTK--LAEPLEIVFESH--ORPPNMTLTCVFDVTKGT-TGDWSSEGCs-TEVR 238
D 852 AAINKEFSNKVYLADPVVTVKHKQSEENFNPCSFWSYKRTMTGYWSTQGCRLLTN 911
QY 239 PEGTVCCDHLTFALLLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFRL 294
D 912 KTHHTCSNHLTNFAVLMAHVEVKHSDAVHDLDDVITWVGILLSVCLLICITFCFF- 970
QY 295 LSRRFKSEDAPKIHVALGSLFLNLAFVNVGSGSGSDAACWARGAVHYFLLCFT 354
D 971 -----RGLQSDRNTIHKNCISLFAELLFL--IGINRTDQPIACAVFAALLHFFFLAAET 1024
QY 355 WMGLEAFHLYLLAVRVNTYEGH--YFLKLSLVGWLPLAMVIGTGSAN--SYGLYTIRD 410
D 1025 WMFLEGVQLYIMLVFESEHSRRKYFY---LVGYMPALIVAVSAVDYRSYG-----1075
QY 411 RENRTSLELCWFREGTMYALYITVHGFLITF-----LFGMVVLALVVMKIFTLR 462
D 1076 -----TDKVCWLRLDT-----YFIWSFGPATLIIMLVIFGLIALYKMF---H 1116
QY 463 ATAV--KERG--KNRKK-VLTLLGLSLVGVTV--GLAIFTPLGLSTVYIFALFNSLOGV 515
D 1117 HTAILKPESGCLDNKISWVIGAILLCLLGLTWAFGLMYINESTVIMAYLFTIFNSLOGM 1176
QY 516 FICCWFTIL 524
D 1177 FIFIFHCVL 1118

RESULT 13
O97818 PRELIMINARY; PRT; 1503 AA.
AC O97818;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE LATROPHILIN 3 SPLICE VARIANT ABAB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148828; PubMed=10025961;
RA Matsushita H., Lelianaova V.G., Ushkaryov Y.A.;
RT "The latrophilin family: multiply spliced G protein-coupled receptors
RT with differential tissue distribution."
RL FBS Lett. 443:348-352(1999).
DR EMBL; AF111085; AAD05321.1;
DR InterPro; IPR000922; Gal.lectin.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; horrm_receptor.
DR InterPro; IPR003334; Latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm.2; 1.
DR Pfam; PF02140; Gal.lectin; 1.
DR Pfam; PF01925; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02354; Latrophilin; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
```

```
ProDom; PD005612; Gal.lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS50228; SUBL_LECTIN; 1.
SQ SEQUENCE 1503 AA; 167991 MW; F4C8345D7F988218 CRC64;

Query Match 13.3%; Score 383.5; DB 6; Length 1503;
Best Local Similarity 27.6%; Pred. No. 1.7e-24;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

QY 77 HLMKEGLTKVNTPELKALVQLNST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
D 683 NLLKTDIVRE-NTNIOLEVARLSTEGNEDLKFPENTGHSTIQLSANTLKQN-----735
QY 131 VRLPKSLFRSLPGNRSVRLAVTIL-DIGPGTLFKGPRGLG-----DGSGLVNLRLVG 183
D 736 -----GRNGEIRVAFVLYNNGPYLSTENASMKLGTEAMSTNHSVIVNSPVIT 783
QY 184 LSVGQMHVTK--LAEPLEIVFESH--ORPPNMTLTCVFDVTKGT-TGDWSSEGCs-TEVR 238
D 784 AAINKEFSNKVYLADPVVTVKHKQSEENFNPCSFWSYKRTMTGYWSTQGCRLLTN 843
QY 239 PEGTVCCDHLTFALLLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFRL 294
D 844 KTHHTCSNHLTNFAVLMAHVEVKHSDAVHDLDDVITWVGILLSVCLLICITFCFF- 902
QY 295 LSRRFKSEDAPKIHVALGSLFLNLAFVNVGSGSGSDAACWARGAVHYFLLCFT 354
D 903 -----RGLQSDRNTIHKNCISLFAELLFL--IGINRTDQPIACAVFAALLHFFFLAAET 956
QY 355 WMGLEAFHLYLLAVRVNTYEGH--YFLKLSLVGWLPLAMVIGTGSAN--SYGLYTIRD 410
D 957 WMFLEGVQLYIMLVFESEHSRRKYFY---LVGYMPALIVAVSAVDYRSYG-----1007
QY 411 RENRTSLELCWFREGTMYALYITVHGFLITF-----LFGMVVLALVVMKIFTLR 462
D 1008 -----TDKVCWLRLDT-----YFIWSFGPATLIIMLVIFGLIALYKMF---H 1048
QY 463 ATAV--KERG--KNRKK-VLTLLGLSLVGVTV--GLAIFTPLGLSTVYIFALFNSLOGV 515
D 1049 HTAILKPESGCLDNKISWVIGAILLCLLGLTWAFGLMYINESTVIMAYLFTIFNSLOGM 1108
QY 516 FICCWFTIL 524
D 1109 FIFIFHCVL 1117
```

```
RESULT 14
O88927 PRELIMINARY; PRT; 1527 AA.
AC O88927; O88924; O88928; O88929; O88926; O88925;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE CALCIUM-INDEPENDENT ALPHA-LATROTOXIN RECEPTOR 3 PRECURSOR (LATROPHILIN
DE 3) (LRP3) (CIRL) (CL3) (LPH3).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047651; PubMed=9830014;
RA Sugita S., Ichtchenko K., Khvotchev M., Suedhof T.C.;
RT "alpha-Latrotoxin receptor CIRL/Latrophilin 1 (CL1) defines an unusual
RT family of ubiquitous G-protein-linked receptors. G-protein coupling
RT not required for triggering exocytosis."
RL J. Biol. Chem. 273:32715-32724(1998).
CC -!- FUNCTION: PERFORMS A GENERAL AND UBIQUITOUS FUNCTION AS G-PROTEIN-
```


DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR003334; latrophilin.
DR InterPro; IPR003112; Olfac_like.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF02140; Gal_Lectin; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF02354; Latrophilin; 1.
DR Pfam; PF02191; OLF; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRODom; PD005612; Gal_lectin; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00284; OLF; 1.
DR PROSITE; PS02027; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS02061; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS02028; SUEL_LECTIN; 1.
SQ SEQUENCE 1571 AA; 174901 MW; 2185D5ADCB5F8607 CRC64;

Query Match 13.3%; Score 383.5; DB 6; Length 1571;
Best Local Similarity 27.6%; Pred. No. 1.8e-24;
Matches 135; Conservative 77; Mismatches 182; Indels 95; Gaps 25;

Qy 77 HLMKEGLTKYNTFFKALVONLST-NTAEDFYFSLEPS-----QVPRQVMKDEKPPDR 130
Db 751 NLLKTDIVRE-NTNIOLEVARLSTENLEDKPEPTNGHSTIQLSANTLKON----- 803

Qy 131 VRLPKSLFRSLPGNRVVRVLAFTIL-DIGPGLTKPRLGLG-----DGSGVLNNRLVG 183
Db 804 -----GRNGELRVAFVLYNLLGPLYLSTENASMKLGTEAMSTNHSVIYNSPVIT 851

Qy 184 LSVQGMHVK--LAEPLEIVFESH-QRPNNMTLCVFDVTKGT-TGDSSEGCST-EYR 238
Db 852 AAINKFESKVLADPVVFTVKHKQSEENFNPCSEWSYKRTMTGYWSTQGCRLTTN 911

Qy 239 PEGTVCCDHLTFALLLR--PTLDQSTVH--ILTRISQAGCGVSMIFLAFTIILYAFLR 294
Db 912 KTHITCSNHLTNFAVLMAHVEVKHSDAVHDLDDVITWYGILLVCLLCICITFCFF- 970

Qy 295 LSRRFRKSEDAPKTHVALGSLFLLNLAFLVNNGSGSGSDAACWARGAVFHYFLLCFT 354
Db 971 ----RGLQSDRNTHKNLCLSLFAVELLEL--IGINRTDQPIACAVFAALLHFFFLAFT 1024

Qy 355 WMGLEAFHYLLAVRVNTYFGH--YFKLSLVGWGLPALMVIGTGSAN--SYGLYTIRD 410
Db 1025 WMFLEGVQLXIMLVEVFESEHSRRKRYF---LVGYGMPALIVAVSAADVRSYG----- 1075

Qy 411 RENRTSLCLCFREGTTHYALYITVHGYPFLITF-----LFGMVVLALVWVKIFTLSR 462
Db 1076 -----TDKVCWRLDIT-----YFIWSFIGPATLIIMLVIFLGIYALYKMF---H 1116

Qy 463 ATAV--KERK--KNRKK-VLTLLGLSLGVGTW--GLAIFTPLGLSTVYIFALFNSLOGV 515
Db 1117 HTAILKPESCLDNKISWVIGATALLCLGLTWFAGFLMYNESTVIMAYLFTIFNSLOGM 1176

Qy 516 FICWFTIL 524
Db 1177 FIFIFHCVL 1185

Search completed: June 11, 2002, 22:32:13
Job time: 457 sec